NIHR | National Institute for Health and Care Research

Health and Social Care Delivery Research

Check for updates

Research Article

Implementing artificial intelligence in chest diagnostics for lung disease: A mixed-methods evaluation

Angus IG Ramsay,[1*] Kevin Herbert,[2] Rachel Lawrence,[1] Chris Sherlaw-Johnson,[3] Stuti Bagri,[3] Nadia Crellin,[3] Emma Dodsworth,[3] Holly Elphinstone,[1] Amanda Halliday,[4] Joanne Lloyd,[4] Efthalia Massou,[2] Raj Mehta,[4] Stephen Morris,[2] Pei Li Ng,[1] Holly Walton[1] and Naomi J Fulop[1]

[1]Department of Behavioural Science and Health, Institute of Epidemiology and Healthcare, University College London, London, UK
[2]Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK
[3]Research and Policy, Nuffield Trust, London, UK
[4]Public Contributor, University College London, London, UK

[*]**Corresponding author** angus.ramsay@ucl.ac.uk

Published March 2026
DOI: 10.3310/GJAR2722

Abstract

Background: Artificial intelligence tools simulate aspects of human intelligence. Policy and research highlight artificial intelligence's potential to support delivery of radiology pathways. In 2023, National Health Service (NHS) England invested £21M to deploy artificial intelligence diagnostic tools for chest X-ray and chest computed tomography in 66 NHS Trusts. Little is known about how artificial intelligence tools are implemented in practice, staff experience of these tools, and their effectiveness and cost.

Objectives:
- Evaluate evidence on artificial intelligence tools within radiology internationally.
- Evaluate implementation of artificial intelligence for chest diagnostics in England.
- Investigate how effectiveness and cost-effectiveness of artificial intelligence for chest diagnostics can be measured.

Methods: Ten-month mixed-methods study (rapid scoping review and empirical study comprising staff interviews, observations and documentary analysis). Findings were analysed using rapid assessment procedures, drawing on qualitative, quantitative and health economic approaches. Our evaluation was also designed to inform phase 2 of our study.

Findings: The review included 114 articles on artificial intelligence use in radiology, internationally. Empirical work included 51 staff interviews, 57 observations and 166 documents from 10/11 of the networks and 6/66 trusts implementing artificial intelligence tools.

Our review found evidence gaps, including real-world implementation of artificial intelligence tools; patient and carer experiences; impact on inequalities, sustainability and wider systems; and cost-effectiveness.

Artificial intelligence for chest diagnostics was implemented in various ways, with different aims, pathways and approaches. Implementation takes time due to multiple activities related to planning, procurement, preparation for deployment, monitoring and evaluation. These tasks – technical and social – required time and resource, including a wide range of stakeholders and expertise. As of November 2024, 24/66 trusts had implemented artificial intelligence tools in practice. Implementation barriers included time, resources, challenges navigating processes and adapting these to local contexts. Facilitators included stakeholder engagement and support. Network and trust ability to evaluate service impact was influenced by factors such as data availability, data linkage, resources and capacity. Factors varied across implementation stages.

Our findings indicated multiple data sources that may support measurement of effectiveness and cost-effectiveness within the English National Health Service. However, limitations to data availability need to be addressed.

This article should be referenced as follows:
Ramsay AIG, Herbert K, Lawrence R, Sherlaw-Johnson C, Bagri S, Crellin N, *et al.* Implementing artificial intelligence in chest diagnostics for lung disease: A mixed-methods evaluation [published online ahead of print March 18 2026]. *Health Soc Care Deliv Res* 2026. https://doi.org/10.3310/GJAR2722

Limitations: Our rapid timeline meant we could not interview patients, carers, and several staff groups at trust and national levels. Delayed deployment meant we could not study implementation in practice or staff experiences.
Future work: Should address real-world implementation, adaptation, and sustainability of artificial intelligence tools; impact of artificial intelligence on care, outcomes, and cost-effectiveness; and staff, patient, and carer experiences of artificial intelligence in practice. We will study these issues in phase 2 of our evaluation.
Conclusions: Artificial intelligence tools may support effective, efficient chest diagnostics services. However, several factors should be considered when implementing and monitoring artificial intelligence tools. Implementation and monitoring may be improved through allowing sufficient time for procurement and preparation for deployment or extending capacity to speed completion of these tasks, early and ongoing stakeholder engagement, sufficient resourcing, dedicated expertise and clinical champions, simplifying governance processes, and improving data capacity. Parallels with learning from implementing other innovations suggests that artificial intelligence tools may not offer straightforward solutions anticipated by services and policy-makers.
Funding: This article presents independent research funded by the National Institute for Health and Care Research (NIHR) Health and Social Care Delivery Research programme as award number NIHR167339.
A plain language summary of this research article is available on the NIHR Journals Library Website https://doi.org/10.3310/GJAR2722.

Overview

What was already known?

- Artificial intelligence (AI) has the potential to support care provision within radiology pathways.
- However, there are some evidence gaps, including real-world implementation, patient/carer experience, impact on inequalities, sustainability and the wider hospital system, and cost-effectiveness.

What this study adds?

- Implementing AI in practice for chest diagnostics involves a complex set of processes, social as well as technical, and therefore takes time and resources. Parallels with other innovations (digital and otherwise) suggest that AI tools may not represent as straightforward a solution to service challenges as anticipated by policy-makers, requiring similar considerations as other innovations.
- Many factors influence the implementation of AI tools for chest diagnostics, including stakeholder engagement, challenges adapting AI to local contexts, availability of champions/expertise, and time and resources.
- For monitoring and evaluation, several key evidence gaps have been identified, including evaluations of real-world AI implementation at various stages, data infrastructure and capacity, and consideration of evaluation plans more broadly.
- Although most suggested evaluation metrics come from existing sources, a few have been derived from our own engagement during the project.
- We provide recommendations on how implementation and monitoring of AI tools for chest diagnostics can be improved.

Background

Artificial intelligence has been defined as advanced technology that can simulate complex tasks associated with human intelligence.[1–4] In recent years, there has been a rapid policy movement towards the implementation and expansion of AI within the English National Health Service (NHS).[5–12] For example, initiatives such as the AI deployment platform,[13] and AI in health and care award.[14] Healthcare policy has indicated the potential benefits of AI for a range of purposes (e.g. administrative support, transforming care pathways, patient triage, optimising staff time).[9] This may be particularly important in the UK, where the NHS faces workforce shortages and demand outweighs resources.[15]

One area in which AI has been thought to demonstrate significant potential is radiology and other imaging services, whereby AI can be used alongside clinicians (as a second reader) to aid early detection of a range of conditions; including cancer (e.g. by assisting with diagnostic reporting),[4,16] stroke[17] and eye conditions.[18] These benefits may help address large backlogs in radiology caused by workforce shortages and their impact on the delivery of safe and effective care.[15]

Findings from previous studies indicate potential benefits of AI in radiology regarding improving detection accuracy,[19] reducing and preventing errors,[19] improving efficiency,[19] supporting clinical decision-making[4,20] and reducing workforce burden.[4] However, there is mixed evidence, with other reviews reporting inconsistencies and variation in diagnostic accuracy,[21] bias in AI technologies[1,2] and the need for infrastructure and workforce capacity to facilitate implementation.[1,3] Finally, previous research indicates that patients and the public may have concerns about the use of AI (e.g. privacy, and confidence in AI).[16,22–24] Previous research highlights limitations in the current evidence base (e.g. sample size, retrospective studies, little evidence on its use within diagnostic pathways).[1] The National Institute for Health and Care Excellence (NICE) has also described where there are evidence gaps for chest diagnostics: for example, gaps on time saving and resource use, adverse effects, performance in different patient groups, ease of use and perceived impacts.[5] Therefore, there may be insufficient evidence to determine the safety, effectiveness and cost-effectiveness of AI tools within chest diagnostics.

In June 2023, NHS England announced the Artificial Intelligence Diagnostic Fund (AIDF), investing £21M to accelerate deployment and implementation of AI diagnostic tools for chest X-rays (CXRs) and chest computed tomography (CT) scans to improve the diagnosis of lung cancer (LC), and to help address current unmet need for faster CXR reporting more generally.[25]

While research suggests AI diagnostic tools may support and improve detection of LC and other chest conditions, little is known about how these tools can be and are being used in real-world settings,[5] experiences and perceptions of implementing AI diagnostic tools, cost and impact on outcomes, (in)equalities and the wider healthcare system.[5]

Aim and research questions

This study therefore aimed to evaluate early deployment and implementation of AI for chest diagnostics as part of the AIDF. This study was phase 1 of a two-phase project. We also aimed to explore ways in which phase 2 and other future evaluations might be conducted. Research questions (RQs) are highlighted in *Table 1*.

Methods

Evaluation design

This rapid study was conducted over a 10-month period (February–November 2024), combining a rapid scoping review and a mixed-methods empirical study (see *Table 1*).

Detailed methods for the review and workshops are provided elsewhere[26] (see *Report Supplementary Material 1*). The review was registered on PROSPERO (CRD42024537518) and followed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis statement.[27]

An empirical study was conducted comprising semi-structured interviews, observations, and documentary analysis.

TABLE 1 Research question and summary of methods used to address each one

Research question	Methods used to answer RQ?			
	Interviews	Observations	Documentary analysis	Rapid scoping review
1. How were AI tools for chest diagnostics procured, deployed, and implemented at network and trust level?	Yes	Yes	Yes	No
2. Which factors influenced implementation at network and trust levels (including contextual factors and implications for EDI?)	Yes	Yes	Yes	Yes[a]
3. What were stakeholder experiences (staff and AI suppliers) of the use of AI in chest diagnostics and associated care pathways?	Yes[b]	Yes	Yes	Yes[a]
4. How can patient and public perceptions and experiences of using AI diagnostic tools in clinical practice be evaluated?	No	No	No	Yes
5. How can services best measure the impact of AI deployment on patients and the clinical pathway (including implications for safety and health inequalities)?	Yes	Yes	Yes	Yes[a]
6. What were the key cost components of AI tools for chest diagnostic that are necessary for an economic evaluation?	Yes	Yes	Yes	Yes[a]

EDI, equality, diversity and inclusion.
a AI diagnostics in radiology more broadly.
b Due to delays in implementation, it was not possible to study stakeholder experiences of using AI in clinical practice in this phase of the evaluation.

Setting

This study took place in England, in 11 NHS networks and 66 trusts that were planning to implement AI chest diagnostic tools as part of the AI Diagnostic Fund.

Theoretical framework

Our evaluation was guided by the Non-adoption, Abandonment and challenges to the Scale-up, Spread and Sustainability (NASSS) of technologies for health and social care framework.[28] This framework was used as it focuses on the implementation of technological innovations and therefore highly relevant to analysing the use of AI in practice.

Sample and recruitment

Rapid scoping review[26]

Studies eligible for inclusion (1) focused on AI being used to support diagnostics in radiology imaging, (2) reported on empirical findings related to implementation, stakeholder experiences and/or perceptions, quantitative or cost outcomes and (3) were published between 2020 and 2024 (present). The search was conducted in April 2024.[26]

Additionally, members of the public with experience of diagnostics, and healthcare professionals working in radiology or collaboratively with radiology teams were recruited to participate in stakeholder workshops to discuss review findings (see *Report Supplementary Material 1*).

Empirical study

To explore experiences across the AIDF, we recruited 10/11 diagnostic imaging networks who were successful in their applications for funds from the AIDF. To explore experiences in more depth at a local level, we recruited six trusts from three networks as case study sites (*Figure 1*).

Sites were recruited through national AIDF meetings and expressions of interest. Sites were sampled using a range of characteristics: diagnostic modality (chest CT or X-ray), AI supplier, geographical location, and stage of implementation.

For interviews, we aimed to recruit one to two staff per network and three to four staff per trust. Eligible participants at both network and trust levels were invited to take part in an interview via e-mail. They were sent an information sheet and consent form and given at least 48 hours to consider participation.

We asked networks and trusts to highlight key meetings that we should observe and to provide key documents relating to AI procurement, deployment and

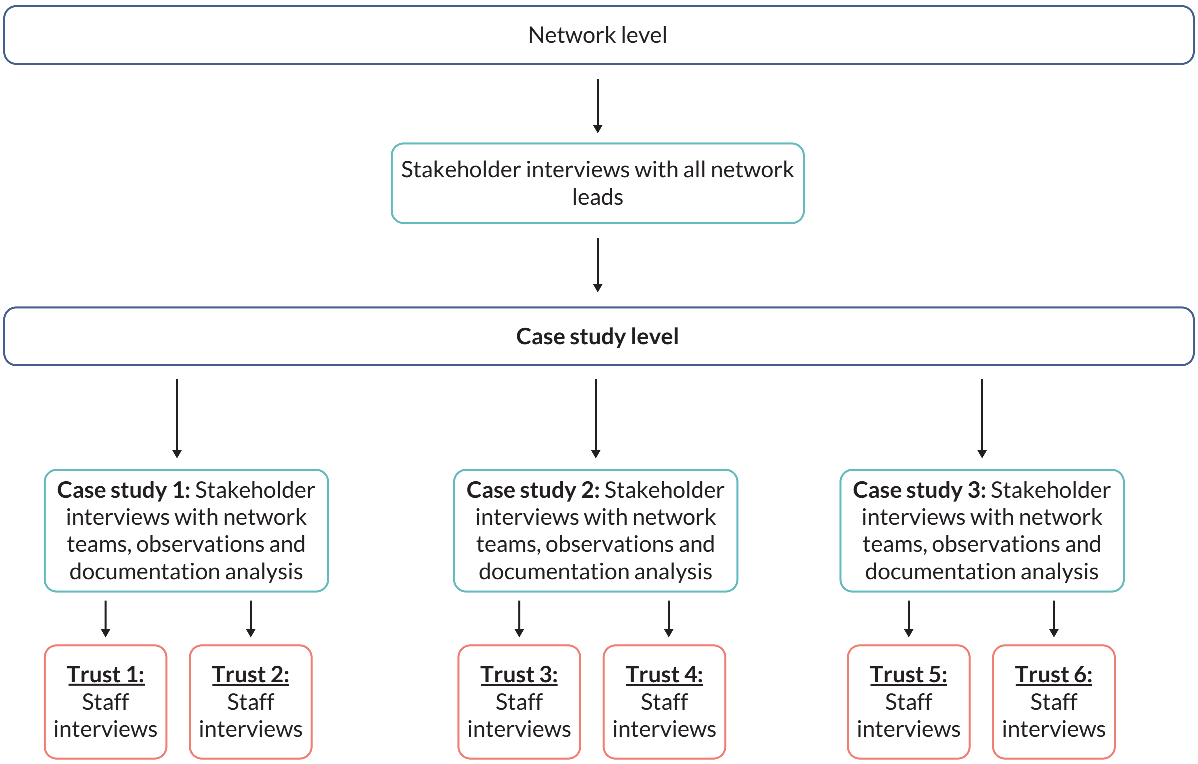

FIGURE 1 Summary of network-level and trust-level (in-depth case study) data collection approaches.

implementation (e.g. specifications, invitations to tender, procurement reports, deployment plans). Meeting chairs provided written consent for observations to take place.

Measures

We developed topic guides for the review stakeholder workshops and empirical study interviews (see *Report Supplementary Material 2*). Topic guides for interviews were informed by the NASSS framework,[28] previous literature,[16,29–31] and discussions within the research team and with wider stakeholders.

Procedure

Rapid scoping review

The review process involved the following key steps: searching the evidence, screening, data extraction and data synthesis.[32] Studies were screened in three phases using the inclusion/exclusion criteria: (1) title, (2) abstracts/summaries and (3) full text. A data extraction tool was developed and used to extract data. Two stakeholder workshops were conducted to discuss review findings. Participants were presented with preliminary findings for discussion within the workshops. Workshops were recorded and transcribed (see *Report Supplementary Material 1*).

Empirical study

We conducted semistructured interviews and meeting observations. Interviews were conducted via Microsoft Teams, audio-recorded and professionally transcribed verbatim by a professional transcription service. Observations of key meetings were conducted online by the research team. During these meetings, researchers took notes outlining the key points, which were included in the analysis.

Data analysis

Rapid scoping review

To answer RQs 4–6, review and workshop findings were analysed using narrative synthesis[33] (see *Report Supplementary Material 1* for detailed description of narrative synthesis).[26] Workshop findings were analysed using a thematic inductive approach[34] structured around initial review findings (see *Report Supplementary Material 1*).[26]

Empirical study

We conducted data collection and analysis in parallel. We used rapid assessment procedure (RAP) sheets, which are templates that can be used to facilitate data analysis.[35] RAP sheets were developed for each network site and the categories included were primarily based on the topic guides and informed by the NASSS framework,[28] but also findings from meeting observations and relevant documents. Notes from interviews and meeting observations were added to the RAP sheets throughout data collection, which were summarised to develop initial themes and subthemes.

To answer RQs 1–3, a coding framework (based on initial themes and sub-themes developed from the RAP sheet analysis) was developed. Researchers coded the transcripts line by line to finalise themes and subthemes, supported by illustrative quotes. We drew on documentary evidence to establish detail on priorities for network procurement, scoring and selection processes, and local implementation of AI tools.

To answer RQ5, we analysed interview findings (together with discussions with sites and suppliers, published literature and wider stakeholder engagement) to develop a list of relevant metrics for evaluating deployments. Additionally, each trust involved in the AIDF was mandated to submit data on the perceived impact of AI deployment based on a template provided by NHSE termed 'benefits metrics'. These were then collated on a network level and submitted to the FutureNHS platform. Baseline (i.e. pre-deployment) data were available during this phase of our study. We assessed the data reported by all networks that made a submission (8 of 11), noting general observations, any inconsistencies in reporting, and fields that were left incomplete. Data on a select list of key metrics were compiled from each of the trusts to provide a comprehensive spreadsheet on all trusts/networks.

To address RQ6, a mapping of diagnostic imaging within the LC care pathway was developed, based initially on the National Optimal Lung Cancer Pathway,[36] with site-specific details extracted from stakeholder interviews. Interview data (and additional sources signposted in interviews) were analysed to inform key costs, resource use and impact metrics. The availability, format and scope of data relevant to the economic study were evaluated in consultation with the quantitative team.

Integration

The qualitative, quantitative and health economic teams co-led interviews and populated the RAP sheets.[35] Emerging findings were shared regularly with the whole team to facilitate interpretation across the workstreams and to ensure that findings from each workstream were complementary in answering the RQs. Further, workstream teams collaborated to present findings in themes structured around the RQs (e.g. qualitative and health economic teams on resource implications of procurement and deployment, and quantitative and health economic teams on data availability).

Patient and public involvement and engagement

The Rapid Service Evaluation Team (RSET) patient and public involvement and engagement (PPIE) co-leads and three public contributors were part of the project team. They attended weekly project meetings and contributed to decision-making on study design, material development, interpretation of findings, dissemination planning and write-up. We also conducted 2 workshops with 16 members of the public with relevant lived experience to gather broader perspectives on the study design and findings. Changes made based on PPIE feedback included adding key questions to the interview topic guide to explore patient safety aspects (see *Report Supplementary Material 3*).

Equality, diversity and inclusion

We considered equality, diversity and inclusion (EDI) throughout the evaluation process; facilitated by the development and use of an EDI checklist. Using this checklist, we were able to consider EDI throughout the project. For example, by including inequalities in the topic guide questions and RAP sheet headings during analysis (see *Report Supplementary Material 4*).

Ethical approval

Ethical approval for this study was provided by University College London (UCL) ethics committee (27037/001) on 18 March 2024. NHS Health Research Authority (HRA) approval was not required for this study as it was categorised as a service evaluation (confirmed through HRA decision tool).

Study registration details

The study was registered on Research Registry, ID: researchregistry10142.

Results

Summary of data

a. Study characteristics (Review)

One hundred and fourteen studies of AI use in radiology for diagnostics were included (see *Appendix 1*, *Table 8* for a summary of study characteristics).

b. Participant characteristics (Empirical study)

Fifty-one interviews with a wide range of staff involved in AI implementation (including five follow-up interviews with network leads; range length of interviews: 20–94 minutes), and 57 observations (e.g. of meetings, safety reviews, and training sessions) were conducted. Additionally, 166 documents (including specifications, scoring summaries, implementation plans) were analysed (*Table 2*).

Findings

How were AI tools for chest diagnostics procured, deployed, and implemented at network and trust level?

AI in diagnostics more broadly (from review)

Six studies evaluated implementation (four process of implementation into clinical pathways, two experiences of implementation). AI was implemented into existing radiology workflows and systems. AI tools were used to assist with diagnostics (and were not autonomous) (see Lawrence *et al.*;[26] *Appendix 1*, *Table 8*). These studies did not study procurement of AI tools.

AI in chest diagnostics (empirical study)

Drivers of implementation Many contextual drivers influenced the implementation of AI for chest diagnostics through the AIDF, including system-, regional- and trust-level drivers. System-level drivers included the need to ease NHS pressures such as significant waiting times and workforce shortages, a growing acknowledgement of the potential for innovation using AI, expanding commercial industry focusing on AI, and variable experience of using AI. Trust-level factors included trust autonomy and local variations in service priorities/pressures, IT infrastructure, clinical pathways, staffing and capacity, and local population (see *Appendix 2*, *Table 9* for example quotes).

Summary of the tools implemented The AI patient pathway is outlined as a summary in *Figure 2*.

Artificial intelligence diagnostic tools for chest diagnostics varied substantially across networks in relation to a range of AI characteristics (including the AI focus, function and chosen supplier) and service pathway characteristics (including referral methods, eligibility criteria, staff responsible for reviewing scans and patient communication) (*Table 3*).

AI characteristics Most networks focused on using AI for CXRs, with a minority focusing on chest CT. Two suppliers (out of all of those that applied) were selected to provide

TABLE 2 Summary of data collected and participant characteristics

		Number of interviews conducted in total (range across networks)	Number of meeting observations in total (range)	Number of documents (range)
Summary of data collected				
Network level (*n* = 10 networks) Initial interviews		31 (range: 2–4) • 26 initial interviews • 5 follow-up interviews	39 (range: 0–15)	166 (range: 2–29)
Trust level (*n* = 6 trusts)		17 (range: 1–5)	18 (range: 2–8)	
Supplier-level interviews		3	N/A	N/A
Total		**51**	**57**	**166**
Participant characteristics – interviews				
Job roles[a]	Network lead	11	N/A	N/A
	Radiologist	9		
	Radiographer	5		
	Project manager	4		
	Trust lead	3		
	PACS staff	3		
	Administrative staff	2		
	Digital diagnostics manager	2		
	Imaging lead	1		
	Clinical systems manager	1		
	Procurement	1		
	Radiology registrar	1		
	Planning consultant	1		
	Data analyst	1		
	Implementation lead	1		
	Supplier/industry	1		

PACS, Picture Archiving and Communication Systems.
a Some interviewees had multiple roles and were interviewed more than once; therefore, the total of interviewees will not add up to 51.

these tools across the 10 participating networks. Many functions of AI were identified, with most networks using AI to identify lung-related conditions such as LC (lung nodules of concern), or chest conditions [such as collapsed lung or tuberculosis (TB)]; however, other functions such as the identification of musculoskeletal fractures and significant safety issues [e.g. misplaced nasogastric (NG) tubes] were also identified. Additionally, a common function of the AI was to identify scans that required prioritisation (processing critical and urgent scans first). Selected functions were linked to local drivers (network and trust need and priorities). For example, prioritisation was key when trusts had significant waiting lists, and identification of misplaced NG tubes were key when trusts had high number of safety events caused by misplaced NG tubes.

Service pathway characteristics Some networks decided to use AI to focus on patients who had been referred from primary care (*n* = 4), or hospital outpatient settings (*n* = 4). Different staffing models were used to review scans (radiologists only, radiologists and radiographers), with the most frequent being a combination of radiologists and radiographers. In terms of communication about AI to patients, there were variations in how this was handled (e.g. informing patients through leaflets or

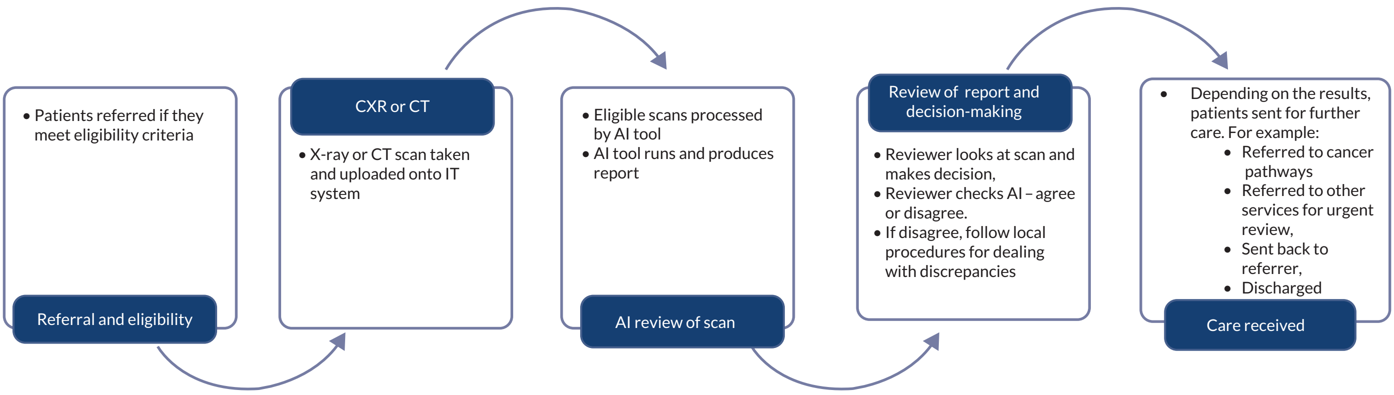

FIGURE 2 Artificial intelligence supported patient pathway for chest diagnostics.

TABLE 3 Variations in implementation

Characteristic		Number of networks (n = 10)[a]
AI characteristics		
Focus of AI	CXR	9
	Chest CT	1
Functions of AI	Identify cancer symptoms or symptoms suspicious of LC	9
	Prioritisation/triage of patients[b]	8
	Identify non cancer symptoms/abnormalities that require care (e.g. sepsis, chest infections, pneumothorax, tuberculosis)	6
	Identify placement/misplacement of nasogastric feeding and breathing tubes	5
	Identify fractures within patients with cancer	3
Supplier	A	6
	B	3
	Still in procurement	1
Service pathway characteristics		
Referral methods	GP referrals	4
	Outpatient referrals	4
	Inpatient	2
	Emergency care	2
	Specialist cancer and cardiac/thoracic hospitals	1
	Any referral mechanism	3
	Not reported	1
Eligibility criteria	All CXRs/CT	3
	> 16 years old	3
	> 18 years old	2
	> 40 years old	1
	All non-urgent (GP/outpatient)	2
	Not reported	1

TABLE 3 Variations in implementation (*continued*)

Characteristic		Number of networks (*n* = 10)[a]
Staffing model	Radiologists only	1
	Radiologists and radiographers	7
	All staff (including radiologists, radiographers, all other doctors and nurse practitioners)	1
	Not reported	1
Communicate with patients about use of AI?	Yes – planning to inform directly (but unclear how/what will tell them)	1
	Yes – planning to inform indirectly (e.g. posters/leaflets up)	3[c]
	Not directly (but sharing details via communications team and media)	1
	Not advertised (but information given if requested by patients)	1
	No	2
	Not reported	2

GP, general practitioner
a For several categories in this table, networks covered more than one characteristic (e.g. multiple functions of AI, referral methods, etc.). Therefore, in a number of cases the numbers do not add up to 10.
b For example critical findings vs. low impact, those who need urgent care to be prioritised or those with normal nodules who can be discharged.
c In one site it varies by trust.

posters, communicating in line with local policies on trust innovations).

Implementation stages and timeline Implementation of AI within networks and trusts required several inter-related complex tasks to occur at various stages (planning for AI implementation, procurement, preparation for deployment and deployment in practice).

NHS England initially anticipated that AI tools would be deployed in practice by December 2023 (procurement October 2023, contracts signed November 2023). However, implementation in practice (and each of the stages) took significantly longer (*Figure 3*). The first trusts deployed the tools in practice in April 2024, and this rose to approximately one third of trusts (24/66) by November 2024. *Figure 3* summarises stages, tasks and timeline.

Procurement Procurement of AI tools was led by imaging networks, with tenders submitted by suppliers who had been approved for the national procurement framework platform. Eligibility for inclusion on the procurement framework was administered by Crown Commercial Services; suppliers were required to submit a range of evidence, including cybersecurity, Medicines and Healthcare products Regulatory Agency compliance, capacity to deploy AI across multiple NHS sites, and cost of AI products. Individual networks worked with local NHS trusts to confirm participation in the programme and recruit trust-level clinical and technical experts to support the procurement process and relevant tasks (e.g. development of specifications and questions for suppliers).

Once applications were received, networks developed procurement panels (comprising network staff, and representation from staff from participating trusts, e.g. radiology, radiography, other clinicians, IT, service managers, finance). Panels undertook an assessment process (which included individual assessments by panel members prior to consensus meetings). Some selection processes were based on documentation alone, whereas others included supplier presentations. Across networks, tender applications were weighted in relation to various criteria (quality, price, social value and cultural fit e.g. environmental sustainability, health inequalities) (*Table 4*). Across networks, quality was generally the highest weighted component. In most networks, selected suppliers scored first for quality but less well on price. In a minority of networks, selected suppliers scored higher for price than quality. EDI was a criterion during procurement; however, this was rarely a differentiating factor in decision-making. Several networks highlighted ongoing discussions with unsuccessful suppliers following decision-making (e.g. regarding validity of procurement processes and decisions). Contracts were then signed with successful suppliers (see *Figure 3* for timings).

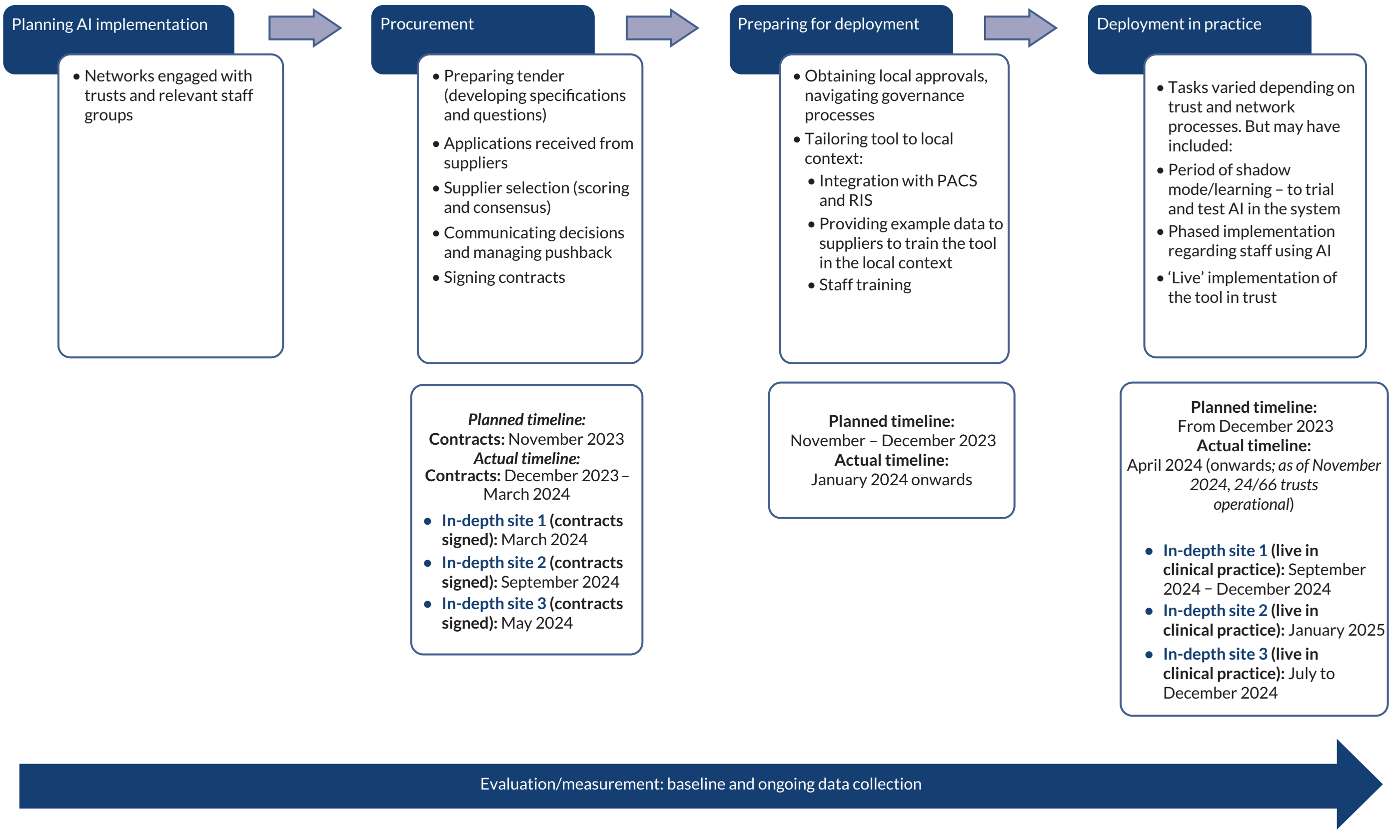

FIGURE 3 Stages, tasks involved and timeline of implementation of AI chest diagnostic tools. PACS, Picture Archiving and Communication System; RIS, Radiology Information System.

TABLE 4 Summary of scoring criteria, weighting and rankings

Criteria used	Number of networks using this criterion (*n* = 10)	Range for weighting of scores across networks (*n* = 10)	Range of rankings for the selected supplier(s) on criteria across the ten networks (*n* = 10)
Quality	10	40–80%	First–third
Price	10	20–50%	First–seventh
Cultural fit	5	10%	First–fifth
Social value	6	10%	First (where reported)
Sustainability	1	10%	First

While procurement happened at network level, national stakeholders (NHS England) produced and shared 'frequently asked questions' (FAQs), a specification template that networks could use (comprising context, tool, operating environment, legal requirements), and held an online workshop. NHS England monitored procurement progress through weekly catch up-meetings with networks.

Preparation for deployment Once suppliers had been selected and contracts signed, networks and trusts began to prepare for deployment. However, this required several processes, including navigating trust governance processes, integrating AI within local trust IT infrastructures and training.

Governance processes included developing and agreeing relevant documentation (e.g. hazard logs, standard operating procedures on how to use AI in different circumstances) and obtaining local clinical safety (CS) and information governance approvals.

Integrating AI *within the trust IT infrastructure* required trusts to work with colleagues from the local trust

Radiology Information System (RIS) and Picture Archiving and Communication System (PACS) to develop and test the AI tools (by running selected scans through AI systems) to ensure they worked (e.g. effective flow of data under different circumstances) and to problem solve when issues arose. Interviewees described differences in how imaging networks approached this. For example, one network standardised the use of AI and workflow integration across all trusts. This was achieved by all participating sites using the same PACS system, which reportedly helped to make a complex process easier. In other networks, trusts used different, rather than standardised systems, and this required further work with individual sites to integrate AI into each radiology workflow.

Training of local staff who would be reporting the scans and looking at the AI reports was required prior to deployment. Training aimed to (1) familiarise staff with how AI reports looked on their systems and how to incorporate these reports into their reporting processes, and (2) educate on the role of AI (decision-support only, to be read after reporter had made their independent assessment). A common approach across networks was to provide training to 'super users' and have them cascade knowledge to colleagues prior to deployment. Teams then needed to monitor who had completed training to ensure that only staff who had received training were using the AI tool (in line with governance priorities). Implemented training monitoring strategies included having specific IT systems that only trained staff could use to access AI, and the use of formal training logs.

Communication with patients and carers was planned variably across networks (see *Table 3*). The majority planned to communicate use of IT passively, for example sharing information if asked by the patient, via posters and leaflets, or via wider trust communications. One Network planned to communicate directly with patients, though were still to decide what information would be shared. Two networks did not plan to communicate information about use of AI. While transparency around AI was highlighted as important by our review focus groups, some interviewees involved in implementation indicated that communication around AI was seen in the same way as communication around other technical changes, for example introduction of a new scanning device.

Deployment/implementation To deploy the AI tools in practice, networks and trusts frequently reported the use of a staged or phased approach. For example, to test that the tool was working within the IT infrastructure and identify issues with information processing, some trusts had a period of 'shadow-mode' whereby AI was launched but not used in live clinical decision-making. Additionally, some services took a phased approach regarding staff able to use the AI; with 'super-users' trialling the AI to process scans and test the system, prior to a wider staff roll-out.

Monitoring/evaluation To support evaluation of the implementation of AI for chest diagnostics, all networks and trusts were required to capture and submit baseline 'benefits metrics' data to the national team (NHS England) on how chest diagnostics systems were operating before AI was introduced. For many networks, this task was undertaken during procurement and preparation for deployment. These would then be compared with similar metrics populated during deployment at intervals of 3, 6 and 12 months.

Which factors influenced implementation at network and trust levels?

AI in diagnostics more broadly (from review)[26]

Several influencing factors were identified from the review, in relation to perceptions and experience covering barriers, facilitators, associated benefits and risks (see *Report Supplementary Material 5*). Barriers included: the high technical demand of AI coupled with a lack of organisational readiness, lack of knowledge and training and a lack of guidance. Facilitators included: transparency (achieved by prioritising staff training and establishing a sense of trust among patients and the public by explaining the use of AI), integrating AI into existing systems and creating AI champions/experts to lead implementation. Perceived risks were staff becoming over-reliant (due to developing algorithmic bias and/or blindness), AI being inaccurate/making errors, deskilling staff, the high cost of AI and ethical risks. The associated benefits were improving reporting times, improving diagnostic accuracy/identifying missed diagnoses and using AI to help with routine tasks.

AI in chest diagnostics (empirical study)

A range of factors influenced implementation of AI tools for chest diagnostics overall, and specifically in relation to the different stages of AI implementation: procurement, preparation for deployment, and deployment (see *Appendix 3*, *Table 10* for summary of barriers and facilitators). These factors commonly reflected technical and social processes central to implementing complex innovations of this kind.

Overall factors influencing implementation

Overall factors influencing the implementation of AI included (1) time, (2) logistics involved in obtaining and processing governance approvals and adapting processes locally, (3) resources and capacity, including project management, (4) service expertise (e.g. in procurement or AI implementation) and championing (e.g. engaging local staff around implementing AI), (5) collaboration – sharing knowledge and learning and (6) monitoring and evaluation (*Table 5* and *Appendix 3*, *Table 10*). All influencing factors are listed in *Table 5*, but two are discussed in more detail below.

Project management Some trusts had a dedicated project manager and interviewees reported that this was an invaluable resource to coordinate and oversee the project across all key phases: from procurement to preparation for deployment and early implementation. When sites did not have a project manager, interviewees described notable challenges as they took on elements of this role, balanced the AI project with their clinical work and were often required to learn as the project progressed. Such differences highlight the central role of project management to support staff with the integration of AI into clinical settings.

Data and monitoring infrastructure Data collection and evaluation were key parts of the project, although for some networks and trusts this was a key barrier. Interviewees reported that trust staff found the data collection process challenging, in relation to the amount of data they needed to collect, the rapid time frames, limited support/guidance and needing to draw out data from different systems to meet the requirements. However, one in-depth site approached this differently and collected required data at a network level, with trusts conducting local evaluations/audits. The rationale for this approach was to reduce burden at a trust level and provide network-wide learning for ongoing and future AI implementation.

TABLE 5 Summary of social and technical factors influencing implementation (overall)

Factor influencing implementation	Summary (overall)
Time	• **Barriers:** Tight timeline for implementation made it difficult to engage staff stakeholders and understand the service environment and the use of AI. Time taken was dependent on scale, complexity of the AI and the number of staff involved
Logistics associated with tailoring, adapting and obtaining approvals at local levels	• **Barriers:** Variations between trusts (in care pathways, availability of specialists and existing IT infrastructure) and information sharing between networks and trusts made it challenging for people leading implementation to tailor, adapt and obtain governance approvals at trust level (e.g. integrating and testing tool with local systems).
Resources/capacity	• **Barriers**: Staff having to undertake this workload on top of their day job due to limited project management resource, and monitoring and integration costs. • **Facilitators:** Dedicated project management was vital in supporting implementation including having a dedicated project manager, regular meetings.
Individual champions and expertise	• **Facilitators:** Implementation was aided by champions who drove local processes amidst competing priorities and who supported engagement with local staff; and services having individuals with expertise (including previous experience with AI and the implementation of AI).
Collaboration and engagement	• **Facilitators:** The collaboration and early engagement between a wide range of staff, including trust-level staff, network-level staff, supplier-level staff and national-level staff.[a] • This collaborative approach also enabled shared learning across networks, with networks meeting regularly to share learning, knowledge and documentation to support progress.
Data and monitoring infrastructure	• **Barriers:** A lack of support for collecting data hindered implementation. In terms of network- and trust-level capacity to collect and provide baseline data required for evaluation, networks and trusts were limited by resources, capacity and time, challenges associated with data availability and data linkage, variability of data systems (PACS and RIS) within networks. • **Facilitators:** Having a plan for evaluation and support with data collection and monitoring, for example from individuals in dedicated data roles, or from Health Innovation Networks, supported implementation as clinical staff did not have to undertake this as part of their role.

a Different groups of individuals played different roles:
- **National stakeholders:** 'honest brokers', providing a framework for implementation, supporting engagement and problem solving.
- **Network-level stakeholders:** engaging relevant stakeholders locally, and supporting planning, project management, monitoring and evaluation.
- **Trust-level stakeholders:** commitment from trust staff to engage with project meetings, governance processes and testing.
- **Suppliers:** supporting planning, document development and training.

Factors influencing different stages of implementation Factors influencing implementation differed across the different stages of implementation (*Table 6* and *Appendix 3*, *Table 10*).

What were stakeholder experiences (staff and AI suppliers) of the use of AI in chest diagnostics and associated care pathways?

AI in diagnostics more broadly (from review)[26]

In most studies included in the review, AI was viewed as a reliable and useable tool in practice. AI was most commonly used as a second reader. However, there were mixed experiences of using AI in practice. Some felt it helped with reducing reading times and improving accuracy but there were concerns about the limited evidence, risk of false positives and reduced efficiency. Evidence was also mixed about the impact of using AI on self-reported confidence. Risks (e.g. over-reliance, inaccuracies, deskilling staff) and benefits (e.g. improving reporting times, diagnostic accuracy) of using AI were identified (see Lawrence *et al.*,[26] *Report Supplementary Material 5*).

AI in chest diagnostics (empirical study)

Due to delays in implementation, we were unable to evaluate staff experiences of using AI in detail. However, staff reported many perceived benefits of using AI tools for chest diagnostics, for staff, patients, and trusts. Perceived benefits for staff included reducing workforce pressure, quicker reporting and diagnoses, staff confidence, extra safety net, standardising practice, and increasing knowledge and skills. Perceived benefits for patients included the potential to reduce inequalities, detect findings, and provide quicker care. Perceived benefits for trusts included cost-savings. Some potential unintended consequences were reported (e.g. litigation concerns, AI guiding decision making for reporter, creating additional workload).

While many staff had not yet used AI in practice, staff reported mixed views on the implementation of AI. Positive views included being interested, excited and keen to use it. Negative views included nervousness and scepticism around the use of AI, queries around accountability for clinical decisions supported by AI, concerns about AI replacing workforce, and fears of AI missing something/misdiagnosing patients. The few staff who had used an AI tool indicated that it was easy to use from the user perspective.

Not all staff had received training by the end of data collection for this study (October 2024). Reasons for this included AI not yet being implemented (training was commonly carried out shortly before going live to maximise knowledge retention), and others due to time barriers to attend training. For those that had attended training, some staff reported that the training did not fill the gaps that they were hoping it would (e.g. wanting to see more hands-on demos of how to use it, together

TABLE 6 Summary of social and technical factors influencing procurement, preparation for deployment and early implementation

Stage of implementation	Example influential factors
Procurement	• Supplier factors (e.g. negotiating challenges) • Procurement scoring factors (e.g. design of procurement questions, presentations of AI vs. what AI can do in local services differing) • Regional and local navigation processes (e.g. balancing decision-making at network and trust levels, established networks supporting engagement) • National support (e.g. national team supporting procurement processes by providing materials and online sessions) • Time (e.g. time constraints in arranging procurement meetings) • Resources and capacity (e.g. having procurement leads) • Collaboration (e.g. support from individuals with expertise from the trust, network and other networks)
Preparation for early deployment	• Time (e.g. to develop and navigate processes) • Logistics associated with obtaining governance approvals and integrating AI into local services (e.g. adaptations for different trusts) • Resources and capacity (e.g. the importance of project management and meetings) • Collaboration (e.g. importance of supplier engagement)
Early implementation[a]	• Logistics involved in obtaining and processing governance approvals and tailoring these to local context (e.g. unclear governance issues, complexity of setting) • Resources and capacity (e.g. NHS pressures, staff having time to attend training or train others) • Local staff characteristics (e.g. training and engagement)

a Limited data on this due to many trusts and networks not yet having implemented in practice.

with cases to practice on). Some staff still did not feel confident following training due to a lack of understanding and knowledge.

How can patient and public perceptions and experiences of using AI diagnostic tools in clinical practice be evaluated?

AI in diagnostics more broadly (from review)[26]

Review findings highlighted that a range of methods have been used to study patient and public perceptions, including workshops, quantitative surveys and qualitative probe interviews. While studies evaluated public perceptions of AI, none evaluated actual patient and carer experience of receiving care facilitated by AI (see Lawrence *et al.*;[26] *Appendix 1*, *Table 8*).

How can services best measure the impact of AI deployment on patients and the clinical pathway (including implications for safety and health inequalities)?

AI in diagnostics more broadly (from review)[26]

Review findings indicated a range of ways in which impact of AI deployment has been measured in radiology diagnostics more broadly, including measuring diagnostic accuracy, reporting times, interpretation times and the effect of reader experience. 15 out of 43 quantitative studies evaluated the impact of AI within a live clinical pathway, while others simulated typical workflow (see *Appendix 1*, *Table 8*). Findings highlighted improvements in diagnostic accuracy and reductions in interpretation time. No studies looked at the impact of AI on wider hospital systems, inequalities, or long-term sustainability of AI (see Lawrence *et al.*[26]).

AI in chest diagnostics (empirical study)

National Institute for Health and Care Excellence (NICE) have outlined a methodology for generating evidence for use of AI in CXRs.[5] This includes appropriate study designs and metrics.

Evaluation metrics Starting from the NICE guidance and further augmenting from a combination of the scoping review,[26] NHS England Benefits Registers and discussion with stakeholders, we have identified a range of potential evaluation metrics. These range from measures of workflow and system efficiency to clinical outcomes. A list is provided in *Appendix 4*, *Table 11*. Most metrics have been already identified by NICE or NHSE, but in some cases we are proposing deeper levels of granularity than they specify. For example, for staff time in reporting one X-ray image, we propose data systems enable capturing this information by category of finding, namely whether normal, suspicious of cancer, and so on. A few metrics came from interviews and discussions with implementation sites, such as same-day follow-up CT scans, and patient trust and anxiety.

These metrics are generally based around the LC diagnostic pathway and designed to evaluate the impact AI has on efficiency measures such as image reporting and turnaround times as well as its impact on rates and timeliness of cancer detection. There will also be outcome measures from other applications of AI we have not listed, for example identifying misplaced NG tubes or rib fractures.

The availability of data Not all of the metrics are measurable within all sites and some may be more relevant to the purpose of the AI being deployed in those sites than others. We discuss the establishing of relevant and feasible metrics below. Some comments on the availability of different data items are provided in *Appendix 4*, *Table 11*.

Case volumes and workflow data and image reporting and turnaround times are recorded in the local RIS or PACS. Although some granularity may be missing (e.g. source of referral and patient comorbidities), it may be possible to disaggregate these data by demographic characteristics and thus assess the impact of AI on inequalities. Further patient clinical details would have to be obtained by linkage to other hospital systems.

Empirical data on the impact on stages of cancer diagnoses require linkage between radiology and cancer registration data. Other useful data linkage would be between radiology and emergency care data for measuring the impact on patients referred by emergency department (ED). A useful way to analyse differences in outcomes would be to follow up cases where the classification of images by human reviewers differs from the AI. However, this can be difficult to measure due to difficulties of capturing these data in a live deployment.

Radiology data and, where available, data from linked data repositories inform the Benefits Registers required by NHSE over the duration of the deployment. These are designed as snapshots before deployment and, subsequently, at different stages of deployment and are provided by the sites themselves. These do measure some of the outcome measures specified above, although they are not all complete, particularly regarding cancer outcomes. See *Appendix 4*, *Table 11* for details.

The feasibility of linking such data varies between trusts and networks as repositories of linked data are not available everywhere. Moreover, there needs to be analytical staff available to support AI evaluations. Indeed, collecting and processing baseline data required for the Benefits Registers was not an easy exercise for several trusts or networks. Factors that influenced the ability of networks and trusts to collect this baseline data are described in *Table 5* and *Appendix 3*, *Table 10*.

Additional considerations for a full evaluation

Establishing robust comparators In some of the interviewed networks, additional system changes occurred concurrently or as part of AI deployment. These included, standardising reporting codes, implementing cloud-based data collection software and other IT projects, reorganising reporting services to prevent reporting times for normal results from being compromised, and the national targeted LC screening programme. Some of these processes were found to be having an impact on reducing backlogs, turnaround times, and the number of abnormal CXR findings even before the AI had been deployed. Moreover, it is likely that pressures on imaging services will increase, irrespective of AI, which may impact on the ability to observe the actual benefits the AI may be having. Isolating AI implementation from other interventions, system changes and underlying trends would be difficult and might significantly compromise any before and after comparisons. There are similar considerations when comparing against sites where there is no AI deployment.

Changing processes in early-stage deployment It is very likely that AI implementation processes, pathway re-organisation and data collection will evolve through the early stages of deployment as networks learn what works well and what does not, and any evaluation must take this into account. Calibration exercises, for example, which determine the thresholds for classifying abnormal AI findings, will be used in the early stages of deployment to control demand for additional imaging services. Over the course of the deployment this will affect changes in volumes of positive CXR findings and follow-up CT scans. Comprehensive understanding of changes in implementation throughout deployment is essential for interpreting results.

Designing a shortlist of metrics There is a variety of main purposes for the AI tool both between and within Networks. These include reducing the backlog and turnaround times, prioritising urgent cases, clearing normal cases more quickly, and/or improving diagnostic accuracy. Given the range of desired outcomes, there is a practical consideration of ensuring a selected shortlist of metrics is relevant, feasible and concise when conducting a full evaluation. They also need to be of sufficient granularity to identify, among other things, different test outcomes and patient characteristics.

What were the key cost components of AI tools for chest diagnostics that are necessary for an economic evaluation?

AI in diagnostics more broadly (from review)[26]

Cost has been evaluated in five studies, using randomised control trials and Markov based models (see Lawrence *et al.*;[26] *Appendix 1*, *Table 8*).

AI in chest diagnostics (empirical study)

The costs required for evaluating AI in chest diagnostics include AI tool costs, staff and non-staff resources and operational costs (see *Appendix 5*, *Table 12* for details of cost, sources of funding, data availability and observations). We identified some limitations when considering the practicality of performing a full economic evaluation of AI in chest diagnostic imaging. While generally available for current practice, many of the key metrics identified will require some time to reach the level of maturity required to be able to draw robust inferences on the impact of AI in chest diagnostic imaging. This issue is further complicated by the fact that many services are operating with case backlogs, which may impact apparent benefits, if/while these are cleared.

Discussion

Key findings

AI tools are used widely in radiology for diagnostics. In implementing AI for chest diagnostics through AIDF, services had different aims for how AI tools will be used and there was variable implementation of these tools in practice. The process of AI implementation was time-consuming, because it involved several stages (procurement, preparation, deployment), each with many associated tasks. Many system- and organisation-level factors influenced implementation of AI tools for chest diagnostics and diagnostics more broadly. For example, facilitators included stakeholder engagement, and champions/expertise. Barriers included time, resources/capacity and challenges adapting AI to local contexts.

Review and empirical study findings highlighted mixed views regarding staff experience of using AI. Findings indicate important research gaps, including real-world implementation, communication with patients and carers, patient/carer experience, impact on inequalities, sustainability and the wider hospital system and cost-effectiveness. A range of data sources may be used to measure the effectiveness and cost-effectiveness of AI tools for chest diagnostics, but there are limitations in the availability of relevant data and complications due to implementation within live pathways (*Figure 4*).

Despite recent policy initiatives for expansion of AI within UK healthcare provision, with so much competition for resources, the cost-effectiveness case needs to be made if AI is to be adopted in diagnostic imaging services. Key to this determination is the real-world clinical efficacy of AI-supported care and its subsequent health benefits, clear evidence for which is in short supply. Value may also be realised more widely, in terms of the workforce (e.g. workload and resourcing) and patient care experience/health-related quality of life – an area that also would also benefit from additional insight. This deficiency in real-world evidence is not unknown for emerging technologies; however, it remains a sizeable obstacle in understanding the true costs and benefits of adopting AI in health care, and the most effective model(s) for implementation. Although this study focuses on the LC care pathway, there is potential for multifunctional AI platforms to be integrated simultaneously across multiple care applications (e.g. LC, fractures, NG tube placement, etc.). While separate evaluations would be required for these, the potential for sharing AI platform costs across a basket of applications may reduce costs and thus improve the likelihood of cost-effectiveness.

How findings relate to previous research

Our empirical study and review findings highlight many potential perceived benefits of AI implementation for chest diagnostics and radiology more broadly; supporting policy[6,8–12] and previous research.[4,15,19,20] While evidence indicates that many trusts already use AI tools in radiology,[15] few research studies have explored implementation in practice. Therefore, study findings extend previous research by addressing some of these

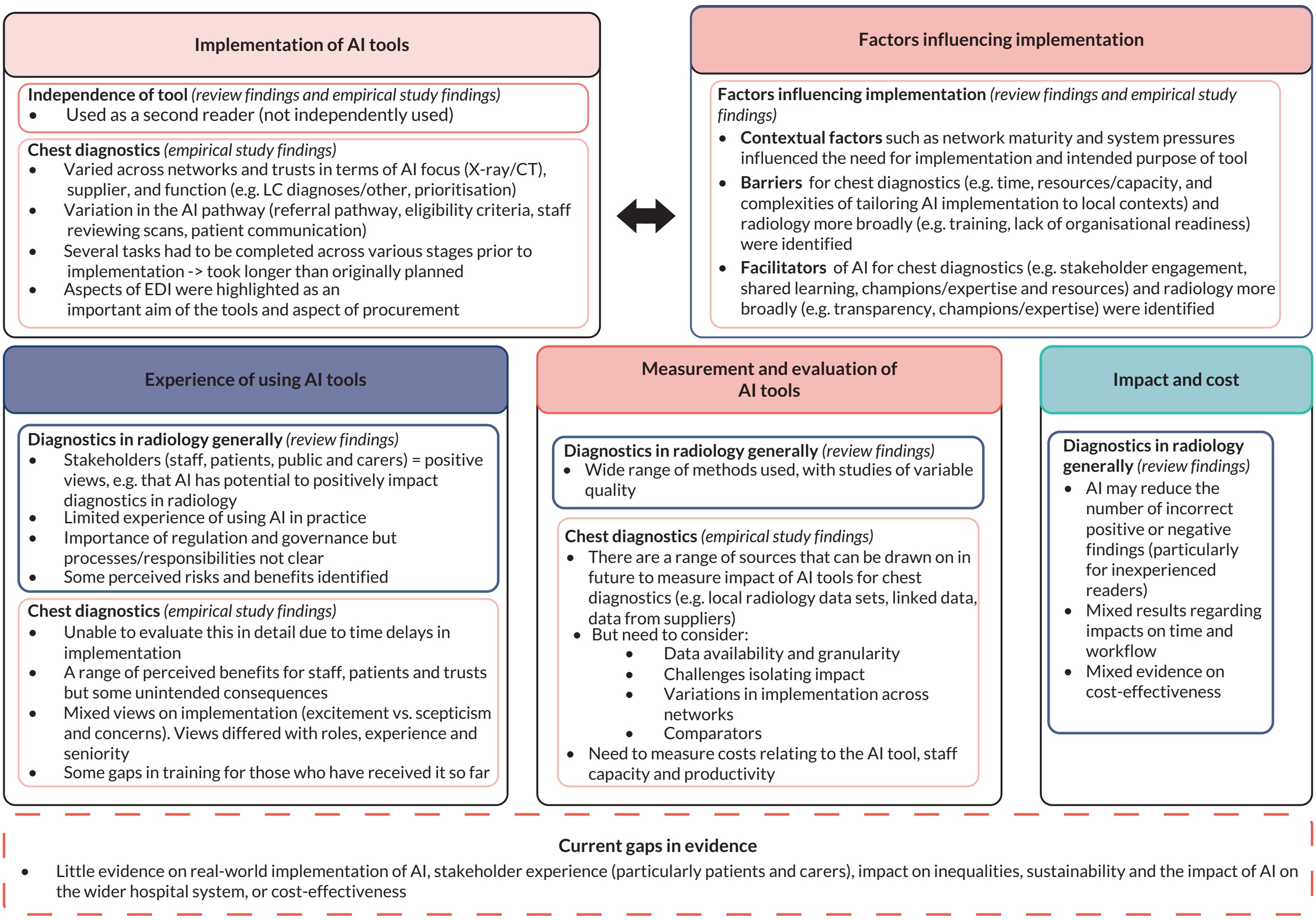

FIGURE 4 Summary of key findings (review and empirical study).

gaps. For example, by providing evidence on how AI has been evaluated to date (in radiology, but often not in live clinical pathways), and providing real-world evidence on AI chest diagnostics in practice, in the stages leading up to implementation including planning, procurement and preparation for deployment. We highlighted the complexity of AI implementation, and the role of many key factors – both social and technical – influencing implementation of AI. This supports and extends previous research which has highlighted key factors influencing AI implementation (e.g. workforce capacity, staff time and training).[1,3] While previous research has outlined the purpose of AI as an assistive tool within radiology systems,[29,37–39] our findings extend evidence by demonstrating variation in use regarding aim, purposes and processes.

Many factors that shaped implementation of AI tools (e.g. system-, organisation- and service-level context, leadership, engagement, technology, data usage and local resourcing) reflect learning from research on implementing other innovations (digital or otherwise).[40–44] The novelty of AI tools suggested a need for greater understanding of and guidance/education on AI tools and their implementation, at system and organisational levels. While these factors clearly influenced procurement and preparation for deployment, many of these factors may well continue to shape usage, experience and impact of AI tools once implemented.

While there is a policy intention to rapidly expand the implementation of AI across the NHS,[6,8–12] our findings are consistent with NICE guidance[5] and previous research,[1,26] which highlight that there are still many unknowns regarding the roll-out of these services into real-world settings. For example, few studies included in the review evaluated effectiveness and cost-effectiveness in real-world clinical pathways. In our empirical study, we extend previous research by identifying the implications of real-world settings for evaluating impact and cost-effectiveness.

Strengths/limitations

A key strength of this evaluation was that it was conducted concurrently with procurement and deployment of AI in chest diagnostics. Also, it included input from various stakeholders to support the planning and delivery of the evaluation, including advice from clinical experts, advisory group members and our PPIE contributors; PPIE took a central role in planning and delivery of this evaluation, shaping the focus of our questions, strengthening our approach to data collection and analysis, and contributing to our interpretation and write-up of findings. Secondly, the evaluation included triangulation across methods (review and empirical study). This enabled an in-depth evaluation of AI in chest diagnostics and radiology more broadly. Additionally, this study sampled 10 out of the 11 implementing networks and included participants from a wide range of roles.

As this evaluation was a rapid evaluation, it was not possible to assess the actual progress of implementation of AI tools for chest diagnostics, due to the delays in deployment of these tools. Therefore, findings on deployment in practice and staff experience of deploying these tools is limited; in particular, staff perspectives on how AI tools may benefit service delivery may have been shaped by optimism and should be revisited once the tools have been fully deployed. In addition, many issues that may influence staff engagement with the tool – including concerns about reliability of the tools (and the potential for automated reading in future), workforce implications (e.g. threats to professions, wider deskilling), and sustainability of AI tools – could not be explored with staff who had worked with the tool in day-to-day clinical practice. Also, while interviewees provided insights on anticipated benefits of AI tool use (e.g. addressing inequalities), it was not possible to assess these benefits in terms of an analysis of the impact of AI tool use in clinical practice.

While we sampled 10 out of 11 networks, it is possible that the trusts that took part reflected some participation bias. That is, we may have recruited sites that were more positive about the programme and associated AI tools, which may in turn have prevented us from capturing a full range of concerns and issues present elsewhere in the AIDF programme sites. Further, our study did not include interviews with patients or carers about their experience of AI, interviews with staff from all relevant roles in trusts (e.g. procurement teams), nor regional or national staff.

Phase 2 of our evaluation is being designed to address several of these limitations, including being able to study full deployment and day-to-day experiences of AI in chest diagnostics and conduct interviews with patients and carers and a wider range of Trust and regional staff.

Implications

Findings from this study have informed the development of several recommendations for implementing and evaluating AI in practice for chest diagnostics. Recommendations have been structured around the AI pathway (*Table 7*).

TABLE 7 Recommendations for implementing and evaluating AI in practice for chest diagnostics

Stage of pathway	Recommendations
Overall	1. There is a need to plan in **sufficient time** for the implementation of AI pathways (allowing for procurement and preparation processes) OR a need to **build in capacity** (e.g. staffing, time) to enable required procurement and preparation steps to be expedited and completed rapidly. *Note. There may be trade-offs with time if services are adapting to the use of AI, and/or laying the foundation for future AI implementation and sustainability.*
	2. There is a need for early and continued **stakeholder engagement** at multiple levels (national, network, trust, IT, supplier, evaluator) to support procurement, preparation and implementation. This requires clear communication and collaboration processes and forums in place to facilitate these.
	3. There is a need for **sufficient resourcing** for services to support procurement, preparation and implementation – including **dedicated project management**, **clinical time** and **evaluation capacity**
	4. **Network-level support is key in supporting trusts** to implement these tools (e.g. with project management, sustaining momentum and facilitating shared learning across network trusts)
	5. Gaps in **patient voice in service development, communications with patients and carers, patient experience** and issues of **equality, diversity and inclusion** must be addressed and considered when developing and implementing AI services
Procurement	6. There is a need for **dedicated expertise** within teams (and access to and awareness of this expertise) to coordinate the process of procurement.
	7. **Communities of practice** (facilitated by national teams) can support local teams to progress with procurement processes by sharing learnings.
	8. There is a need for AI suppliers to consider how **to showcase functionality of** AI **tools** during the procurement process, to mitigate against clinical team uncertainty regarding what AI can achieve.
	9. While NHS Trusts have autonomy in procuring services, there may be value in considering national procurement of tools, as this may free up network capacity to facilitate local deployment
Preparation for deployment	10. Governance processes took time and were challenging for networks and trusts. Therefore, consideration of whether **further support can be provided to complete governance processes**, or whether **governance processes can be simplified** may be helpful. Some possible solutions include: (1) national-level provision of expert informed guidance on who is responsible/how to implement AI/how to mitigate challenges, (2) having clear points of contact, (3) having clear timelines, (4) specific AI resource templates, and (5) having support from suppliers (e.g. involvement in planning meetings etc.).
Staff experience	11. In local services, there is a need for **clinical champions** to drive AI forward, support staff and mitigate staff worries/concerns.
	12. **Education and knowledge gaps** need to be addressed by providing sufficient and consistent training and education to staff using AI. Staff need protected time to attend training themselves, and where relevant to support the training of other staff (e.g. if staff are 'superusers').
	13. **There is a need for open discussions and early engagement with on the ground staff** to understand and address concerns
Monitoring and evaluation	14. Gaps in **evidence** must be addressed to better inform the developing of services, including: (1) evaluations of real-world pathways of AI implementation at various stages, (2) better data infrastructure and capacity within local and national teams (data collection and staff) and (3) consideration of evaluation plans more broadly.

Note
Bold indicates key aspects of recommendations.

Future research

Future research is needed to evaluate real-world implementation, adaptation and sustainability of AI tools, real-world impact of AI on care delivery, outcomes and cost-effectiveness, and to evaluate staff, patient and carer experiences of AI in practice (including ethical considerations around patient information and consent, transparency around algorithms, and implications for inequalities). It is planned that these gaps will be addressed within our phase 2 evaluation if appropriate data are available (e.g. see *How can services best measure the impact of AI deployment on patients and the clinical pathway (including implications for safety and health inequalities)?* and *What were the key cost components of AI tools for chest diagnostics that are necessary for an economic evaluation?*). *Appendix 6*, *Table 13* presents how the research presented here will shape our phase 2 evaluation.

In addition, as noted under *Methods*, our research tools were shaped in part by the NASSS framework: we plan to formally analyse our data using NASSS as the basis for an academic journal article, to enhance the theoretical contributions of this research.

Conclusion

Artificial intelligence has the potential to positively impact diagnostics in radiology, but further evidence is still required. Within chest diagnostics, findings highlighted variable implementation of these tools in practice. The process of AI implementation is time-consuming and complex; many factors influence implementation throughout procurement, preparation for deployment, and deployment. In terms of monitoring and evaluation, findings highlighted several important evidence gaps (e.g. patient/carer experience, impact on inequalities, sustainability, and wider systems within real-world clinical practice and cost-effectiveness). Findings illustrated gaps in data and limitations of data infrastructure that need to be addressed. The influential factors identified in this study suggest that AI tools have much in common with other innovations; therefore, AI tools may not represent as straightforward a solution to service challenges as anticipated by policy-makers, and will require similar considerations and supports as other innovations.

Additional information

CRediT contribution statement

Angus IG Ramsay (https://orcid.org/0000-0002-4446-6916): Conceptualisation, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualisation, Writing – original draft, Writing – reviewing and editing.

Kevin Herbert (https://orcid.org/0009-0008-4354-7811): Conceptualisation, Data curation, Formal analysis, Investigation, Methodology, Resources, Validation, Visualisation, Writing – original draft, Writing – reviewing and editing.

Rachel Lawrence (https://orcid.org/0000-0003-3743-4622): Conceptualisation, Data curation, Formal analysis, Investigation, Methodology, Resources, Validation, Visualisation, Writing – original draft, Writing – reviewing and editing.

Chris Sherlaw-Johnson (https://orcid.org/0000-0002-4851-6060): Conceptualisation, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Supervision, Validation, Visualisation, Writing – reviewing and editing.

Stuti Bagri (https://orcid.org/0009-0005-1267-2503): Conceptualisation, Data curation, Formal analysis, Investigation, Methodology, Resources, Validation, Visualisation, Writing – original draft, Writing – reviewing and editing.

Nadia Crellin (https://orcid.org/0000-0002-9497-5874): Conceptualisation, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Supervision, Validation, Visualisation, Writing – reviewing and editing.

Emma Dodsworth (https://orcid.org/0009-0007-2264-7277): Conceptualisation, Data curation, Formal analysis, Investigation, Methodology, Resources, Supervision, Validation, Visualisation, Writing – reviewing and editing.

Holly Elphinstone (https://orcid.org/0009-0007-3850-2128): Conceptualisation, Methodology, Project administration, Resources, Writing – reviewing and editing.

Amanda Halliday (https://orcid.org/0009-0002-5905-0566): Conceptualisation, Methodology, Writing – reviewing and editing.

Joanne Lloyd (https://orcid.org/0009-0007-7419-1662): Conceptualisation, Methodology, Writing – reviewing and editing.

Efthalia Massou (https://orcid.org/0000-0003-0488-482X): Conceptualisation, Data curation, Formal analysis, Investigation, Methodology, Resources, Validation, Visualisation, Writing – original draft, Writing – reviewing and editing.

Raj Mehta (https://orcid.org/0000-0002-4003-530X): Conceptualisation, Funding acquisition, Methodology, Writing – reviewing and editing.

Stephen Morris (https://orcid.org/0000-0002-5828-3563): Conceptualisation, Funding acquisition, Methodology, Supervision, Writing – reviewing and editing.

Pei Li Ng (https://orcid.org/0000-0001-8411-220X): Conceptualisation, Methodology, Project administration, Resources, Writing – reviewing and editing.

Holly Walton (https://orcid.org/0000-0002-8746-059X): Conceptualisation, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualisation, Writing – original draft, Writing – reviewing and editing.

Naomi J Fulop (https://orcid.org/0000-0001-5306-6140): Conceptualisation, Funding acquisition, Methodology, Supervision, Writing – reviewing and editing.

Acknowledgements

We thank the networks, trusts and participants that have participated in the evaluation, and the NHS England stakeholders who facilitated our work throughout the evaluation. We also thank our clinical advisors (Prof Fergus Gleeson and Dr Tracy O'Regan) and our project advisory group for providing guidance throughout our evaluation, and all stakeholders who contributed to our scoping processes.

Data-sharing statement

The anonymised qualitative data that support the findings of this study are available on request from the corresponding authors. To protect anonymity of interviewees, the data are not publicly available. Further information can be obtained from the corresponding author.

Ethics statement

Ethical approval for this study was provided by UCL ethics committee (27037/001) on 18 March 2024. NHS HRA approval was not required for this study as it was categorised as a service evaluation (confirmed through HRA decision tool).

Information governance statement

University College London, the University of Cambridge and the Nuffield Trust are committed to handling all personal information in line with the UK Data Protection Act (2018) and the General Data Protection Regulation (EU GDPR) 2016/679.

Under the Data Protection legislation, University College London, the University of Cambridge and the Nuffield Trust are the Data Controllers, and you can find out more about how we handle personal data, including how to exercise your individual rights and the contact details for our Data Protection Officer here: www.ucl.ac.uk/data-protection/data-protection-0

Disclosure of interests

Full disclosure of interests: Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at https://doi.org/10.3310/GJAR2722.

Primary conflicts of interest: Dr Angus Ramsay is a trustee at Health Services Research UK. Raj Mehta is Chair of the Board of Trustees of the Middlesex Association for the Blind; Vice-Chair on the Board of Trustees of the Research Institute for Disabled Consumers; Trustee on the Board of the Thomas Pocklington Trust; Non-Executive Director on the Board of Evenbreak; cochair and Director on the Board of Shaping Our Lives. Professor Stephen Morris is currently (2022–present) a member of the Small Business Research Initiative (SBRI) Healthcare panel. His post is funded in part by RAND Europe, a non-profit research organisation. Professor Morris is also Deputy Director of Applied Research Collaboration East of England (NIHR ARC EoE) at Cambridgeshire and Peterborough NHS Foundation Trust. Professor Naomi J Fulop was a member of the UKRI and NIHR College of Experts for COVID-19 Research Funding 2020, a Non-Executive Director at Whittington Health NHS Trust until October 2024 and is a Non-Executive Director at COVID-19 Bereaved Families for Justice UK.

Department of Health and Social Care disclaimer

Study registration

This study is registered on Research Registry, ID: researchregistry10142.

Funding

This article presents independent research funded by the National Institute for Health and Care Research (NIHR) Health and Social Care Delivery Research programme as award number NIHR167339.

This article reports research award *Mixed-method evaluation of implementing artificial intelligence in chest diagnostics for lung disease*. For more information about this research, please view the award page (www.fundingawards.nihr.ac.uk/award/NIHR167339).

About this article

The contractual start date for this research was in November 2023. This article began editorial review in December 2024 and was accepted for publication in August 2025. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The Health and Social Care Delivery Research editors and publisher have tried to ensure the accuracy of the authors' article and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this article.

Copyright

– NIHR Journals Library, and the DOI of the publication must be cited.

Copyright and credit statement

List of supplementary material

Report Supplementary Material 1
Review and workshop methods

Report Supplementary Material 2 Topic guides

Report Supplementary Material 3
Impact of PPIE contributions for this study, structured using the GRIPP2 – Patient and Public Involvement and Engagement (PPIE) Reporting Checklist

Report Supplementary Material 4
Summary of equality, diversity and inclusion (EDI) considerations and EDI checklist

Report Supplementary Material 5
Factors influencing AI adoption (summary of review findings)

Supplementary material can be found on the NIHR Journals Library article page (https://doi.org/10.3310/GJAR2722).

Supplementary material has been provided by the authors to support the article and any files provided at submission will have been seen by peer reviewers, but not extensively reviewed. Any supplementary material provided at a later stage in the process may not have been peer reviewed.

List of abbreviations

AI	artificial intelligence
AIDF	Artificial Intelligence Diagnostic Fund
CS	clinical safety
CT	computed tomography
CXR	chest X-ray
ED	emergency department
EDI	equality, diversity and inclusion
FAQ	frequently asked questions
HRA	Health Research Authority
LC	lung cancer
NASSS	Non-adoption, Abandonment, Scale-up, Spread, Sustainability
NG	nasogastric
NICE	National Institute for Health and Care Excellence
PACS	Picture Archiving and Communication Systems
PPIE	patient and public involvement and engagement
RAP	rapid assessment procedure
RIS	Radiology Information System
RQ	research question
RSET	Rapid Service Evaluation Team
TB	tuberculosis
UCL	University College London

References

1. Gleeson F, Revel MP, Biederer J, Larici AR, Martini K, Frauenfelder T, *et al*. Implementation of artificial intelligence in thoracic imaging-a what, how, and why guide from the European Society of Thoracic Imaging

(ESTI). *Eur Radiol* 2023;**33**:5077–86. https://doi.org/10.1007/s00330-023-09409-2

2. Harwich E, Laycock K. *Thinking on Its Own: AI in the NHS*. Reform Research Trust; 2018. URL: https://reform.uk/wp-content/uploads/2018/11/AI-in-Healthcare-report_WEB.pdf (accessed 16 July 2025).
3. Istasy P, Lee WS, Iansavichene A, Upshur R, Gyawali B, Burkell J, *et al*. The impact of artificial intelligence on health equity in oncology: scoping review. *J Med Internet Res* 2022;**24**:e39748. https://doi.org/10.2196/39748
4. Joy Mathew C, David AM, Joy Mathew CM. Artificial intelligence and its future potential in lung cancer screening. *EXCLI J* 2020;**19**:1552–62. https://doi.org/10.17179/excli2020-3095
5. National Institute for Health and Care Excellence. *Artificial Intelligence-derived Software to Analyse Chest X-rays for Suspected Lung Cancer in Primary Care Referrals: Early Value Assessment*; 2023. URL: www.nice.org.uk/guidance/hte12/resources/artificial-intelligencederived-software-to-analyse-chest-xrays-for-suspected-lung-cancer-in-primary-care-referrals-early-value-assessment-pdf-50261967918277 (accessed 16 July 2025).
6. NHS England. *The NHS Long Term Plan*. 2019. URL: https://web.archive.org/web/20250218094628/https://www.england.nhs.uk/wp-content/uploads/2022/07/nhs-long-term-plan-version-1.2.pdf (accessed 16 July 2025).
7. NHS England. *Artificial Intelligence (AI) and Machine Learning*. 2023. URL: https://webarchive.nationalarchives.gov.uk/ukgwa/20230418155402/https:/www.longtermplan.nhs.uk/publication/nhs-long-term-plan/ (accessed 18 February 2026).
8. Darzi A. *Summary Letter from Lord Darzi to the Secretary of State for Health and Social Care*. 2024. URL: www.gov.uk/government/publications/independent-investigation-of-the-nhs-in-england/summary-letter-from-lord-darzi-to-the-secretary-of-state-for-health-and-social-care (accessed 16 July 2025).
9. NHS Confederation. *AI in Healthcare: Navigating the Noise. A Comprehensive Guide Supporting Healthcare Leaders to Make Sense of AI and Explore the Art of the Possible*. 2024. URL: www.nhsconfed.org/publications/ai-healthcare#:~:text=The%20rapid%20expansion%20of%20artificial,overcome%20a%20range%20of%20challenges (accessed 16 July 2025).
10. NHS England. *Integrated Health and Care*. 2024. URL: www.england.nhs.uk/long-read/integrated-health-and-care/ (accessed 16 July 2025).
11. Darzi A. *Investigation of the National Health Service in England*. 2024. URL: https://assets.publishing.service.gov.uk/media/66f42ae630536cb92748271f/Lord-Darzi-Independent-Investigation-of-the-National-Health-Service-in-England-Updated-25-September.pdf (accessed 16 July 2025).
12. Labour Party. *Wes Streeting Speech at Labour Party Conference 2024*. 2024. URL: https://labour.org.uk/updates/press-releases/wes-streeting-speech-at-labour-party-conference-2024/ (accessed 16 July 2025).
13. NHS England. *AI Deployment Platform*. n.d. URL: https://transform.england.nhs.uk/ai-lab/ai-lab-programmes/ai-in-imaging/ai-deployment-platform/ (accessed 16 July 2025).
14. NHS England. *Artificial Intelligence in Health and Care Award*. n.d. URL: www.england.nhs.uk/aac/what-we-do/how-can-the-aac-help-me/ai-award/ (accessed: 16 July 2025).
15. The Royal College of Radiologists. *Clinical Radiology Workforce Census*. 2023. URL: www.rcr.ac.uk/media/5befglss/rcr-census-clinical-radiology-workforce-census-2023.pdf (accessed 16 July 2025).
16. Zhang Z, Citardi D, Wang D, Genc Y, Shan J, Fan X. Patients' perceptions of using artificial intelligence (AI)-based technology to comprehend radiology imaging data. *Health Informatics J* 2021;**27**:14604582211011215. https://doi.org/10.1177/14604582211011215
17. National Institute for Health and Care Excellence. *Artificial Intelligence (AI)-derived Software to Help Clinical Decision Making in Stroke*. 2024. URL: www.nice.org.uk/guidance/dg57/resources/artificial-intelligence-aiderived-software-to-help-clinical-decision-making-in-stroke-pdf-1053876693445 (accessed 16 July 2025).
18. National Institute for Health and Care Excellence. *AI Technologies for Detecting Diabetic Retinopathy*. 2021. URL: www.nice.org.uk/advice/mib265/resources/ai-technologies-for-detecting-diabetic-retinopathy-pdf-2285965755558853 (accessed 16 July 2025).
19. Liu M, Wu J, Wang N, Zhang X, Bai Y, Guo J, *et al*. The value of artificial intelligence in the diagnosis of lung cancer: a systematic review and meta-analysis. *PLOS ONE* 2023;**18**:e0273445. https://doi.org/10.1371/journal.pone.0273445
20. Chiu HY, Chao HS, Chen YM. Application of artificial intelligence in lung cancer. *Cancers (Basel)* 2022;**14**:1370. https://doi.org/10.3390/cancers14061370
21. Fletcher RR, Nakeshimana A, Olubeko O. Addressing fairness, bias, and appropriate use of artificial intelligence and machine learning in global health. *Front Artif Intell* 2020;**3**:561802. https://doi.org/10.3389/frai.2020.561802

22. Adams SJ, Tang R, Babyn P. Patient perspectives and priorities regarding artificial intelligence in radiology: opportunities for patient-centered radiology. *J Am Coll Radiol* 2020;**17**:1034–6. https://doi.org/10.1016/j.jacr.2020.01.007

23. Fritsch SJ, Blankenheim A, Wahl A, Hetfeld P, Maassen O, Deffge S, *et al*. Attitudes and perception of artificial intelligence in healthcare: a cross-sectional survey among patients. *Digit Health* 2022;**8**. https://doi.org/10.1177/20552076221116772

24. Ibba S, Tancredi C, Fantesini A, Cellina M, Presta R, Montanari R, *et al*. How do patients perceive the AI-radiologists interaction? Results of a survey on 2119 responders. *Eur J Radiol* 2023;**165**:110917. https://doi.org/10.1016/j.ejrad.2023.110917

25. NHS England. *AI Diagnostic Fund*. 2023. URL: https://web.archive.org/web/20250421173708/https://transform.england.nhs.uk/ai-lab/ai-lab-programmes/ai-in-imaging/ai-diagnostic-fund/ (accessed 18 February 2026).

26. Lawrence R, Dodsworth E, Massou E, Sherlaw-Johnson C, Ramsay A, Walton H, *et al*. Artificial Intelligence for diagnostics in radiology practice: a rapid systematic scoping review. *EClinicalMedicine* 2025;**83**:103228. https://doi.org/10.1016/j.eclinm.2025.103228

27. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLOS Med* 2009;**6**:e1000097. https://doi.org/10.1371/journal.pmed.1000097

28. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, *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**:e367. https://doi.org/10.2196/jmir.8775

29. Faric N, Hinder S, Williams R, Ramaesh R, Bernabeu MO, van Beek E, Cresswell K. Early experiences of integrating an artificial intelligence-based diagnostic decision support system into radiology settings: a qualitative study. *Stud Health Technol Inform* 2023;**309**:240–1. https://doi.org/10.3233/SHTI230787

30. Rainey C, O'Regan T, Matthew J, Skelton E, Woznitza N, Chu KY, *et al*. UK reporting radiographers' perceptions of AI in radiographic image interpretation – current perspectives and future developments. *Radiography (Lond)* 2022;**28**:881–8. https://doi.org/10.1016/j.radi.2022.06.006

31. Ng CT, Roslan SNA, Chng YH, Choong DAW, Chong AJL, Tay YX, *et al*. Singapore radiographers' perceptions and expectations of artificial intelligence – a qualitative study. *J Med Imaging Radiat Sci* 2022;**53**:554–63. https://doi.org/10.1016/j.jmir.2022.08.005

32. Tricco AC, Langlois EV, Straus SE. *Rapid Reviews to Strengthen Health Policy Systems: A Practical Guide*. 2017. URL: https://wkc.who.int/resources/publications/i/item/2017-08-10-rapid-reviews-to-strengthen-health-policy-and-systems-a-practical-guide (accessed 16 July 2025).

33. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, *et al*. *Guidance on the Conduct of Narrative Synthesis in Systematic Reviews: Lancaster University*. 2006. URL: www.researchgate.net/publication/233866356_Guidance_on_the_conduct_of_narrative_synthesis_in_systematic_reviews_A_product_from_the_ESRC_Methods_Programme (accessed 16 July 2025).

34. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;**3**:77–101. https://doi.org/10.1191/1478088706qp063oa

35. Vindrola-Padros C, Chisnall G, Cooper S, Dowrick A, Djellouli N, Symmons SM, *et al*. Carrying out rapid qualitative research during a pandemic: emerging lessons from COVID-19. *Qual Health Res* 2020;**30**:2192–204. https://doi.org/10.1177/1049732320951526

36. Roy Castle Lung Cancer Foundation Expert Group. *National Optimal Lung Cancer Pathway (NOLCP) Version 4*. 2024. URL: https://roycastle.org/for-health-care-professionals/clinical-expert-group/ (accessed 16 July 2025).

37. Strohm L, Hehakaya C, Ranschaert ER, Boon WPC, Moors EHM. Implementation of artificial intelligence (AI) applications in radiology: hindering and facilitating factors. *Eur Radiol* 2020;**30**:5525–32. https://doi.org/10.1007/s00330-020-06946-y

38. Kim B, Romeijn S, van Buchem M, Mehrizi MHR, Grootjans W. A holistic approach to implementing artificial intelligence in radiology. *Insights Imaging* 2024;**15**:22. https://doi.org/10.1186/s13244-023-01586-4

39. Wenderott K, Krups J, Luetkens JA, Weigl M. Radiologists' perspectives on the workflow integration of an artificial intelligence-based computer-aided detection system: a qualitative study. *Appl Ergon* 2024;**117**:104243. https://doi.org/10.1016/j.apergo.2024.104243

40. Parmar J, Sacrey LA, Anderson S, Charles L, Dobbs B, McGhan G, *et al*. Facilitators, barriers and considerations for the implementation of healthcare innovation: a qualitative rapid systematic review. *Health Soc Care Community* 2022;**30**:856–68. https://doi.org/10.1111/hsc.13578

41. Santos WJ, Graham ID, Lalonde M, Demery Varin M, Squires JE. The effectiveness of champions in implementing innovations in health care: a systematic review. *Implement Sci Commun* 2022;**3**:80. https://doi.org/10.1186/s43058-022-00315-0

42. Syeed MS, Poudel N, Ngorsuraches S, Veettil SK, Chaiyakunapruk N. Characterizing attributes of innovation of technologies for healthcare: a systematic review. *J Med Econ* 2022;**25**:1158–66. https://doi.org/10.1080/13696998.2022.2140591

43. Milella F, Minelli EA, Strozzi F, Croce D. Change and innovation in healthcare: findings from literature. *ClinicoEcon Outcomes Res* 2021;**13**:395–408. https://doi.org/10.2147/CEOR.S301169

44. Shin HD, Hamovitch E, Gatov E, MacKinnon M, Samawi L, Boateng R, *et al*. The NASSS (Non-Adoption, Abandonment, Scale-Up, Spread and Sustainability) framework use over time: a scoping review. *PLOS Digital Health* 2025;**4**:e0000418. https://doi.org/10.1371/journal.pdig.0000418

Appendix 1

TABLE 8 Review study characteristics

Study characteristics		Overall (*n* = 114)
Study focus[a]	Implementation	6
	Experience	12
	Perceptions	59
	Effectiveness (quantitative studies)	43
	Cost	5
Focus across multiple topics[a]	Implementation, experiences and perceptions	2
	Experiences and perceptions	2
	Implementation and perceptions	1
	Effectiveness and cost-effectiveness	2
	Effectiveness and experience	2

a The total will not always add up to 114 as some papers had multiple areas of focus (as shown above), were conducted across countries, covered different imaging modalities and patient pathways and recruited multiple stakeholder groups.

			Study focus[a]				
	Study characteristics	Overall (*n* = 114)	Implementation (*n* = 6)	Experience (*n* = 12)	Perceptions (*n* = 59)	Effectiveness (quantitative) (*n* = 43)	Cost (*n* = 5)
Location							
	USA	12	0	3	5	5	2
	UK	15	3	1	12	1	1
	Germany	9	0	1	5	3	1
	China	12	0	0	1	11	0
	China and Germany	1	0	0	0	1	0
	Republic of Korea	10	0	3	1	7	0
	Australia	5	0	0	3	2	0
	Saudi Arabia	5	0	0	5	0	0
	The Netherlands	5	3	1	1	2	0
	Spain	3	0	0	3	0	0
	Italy	6	0	1	3	3	0

Study characteristics	Overall (*n* = 114)	Study focus[a] Implementation (*n* = 6)	Experience (*n* = 12)	Perceptions (*n* = 59)	Effectiveness (quantitative) (*n* = 43)	Cost (*n* = 5)
Japan	3	0	0	0	2	1
World-wide	3	0	0	3	0	0
Europe-wide	2	0	1	1	0	0
Africa-wide	2	0	0	2	0	0
United Arab Emirates	2	0	0	2	0	0
France	1	0	0	0	1	0
Nigeria	1	0	0	1	0	0
Middle East and India	1	0	0	1	0	0
India	1	0	0	1	0	0
Ireland	1	0	0	1	0	0
Jordan	1	0	0	1	0	0
Australia and New Zealand	1	0	0	1	0	0
Ghana	1	0	0	1	0	0
Singapore	2	0	0	1	1	0
Nordic countries	1	0	0	1	0	0
Argentina	1	0	1	0	0	0
Malaysia	1	0	0	1	0	0
Vietnam	1	0	0	0	1	0
Egypt	1	0	0	0	1	0
Taiwan (Province of China)	1	0	0	0	1	0
Switzerland	1	0	0	0	1	0
Canada	1	0	0	1	0	0
Unclear	1	0	0	1	0	0
Imaging modalities						
X-ray	16	1	2	1	13	1
CT	14	1	2	2	12	0
MRI	6	1	1	1	5	0
CT or MRI	1	0	0	1	0	0
Colonoscopy	7	0	1	1	5	2
Ultrasound	7	0	0	0	7	0
Mammography	1	0	0	0	1	0
MRI/MRI fusion biopsy	1	0	0	1	0	0
Computed tomography angiography (CTA)	1	0	0	0	0	1

Study characteristics	Overall (*n* = 114)	Study focus[a] Implementation (*n* = 6)	Experience (*n* = 12)	Perceptions (*n* = 59)	Effectiveness (quantitative) (*n* = 43)	Cost (*n* = 5)
Images acquired during endoscopic procedures	1	0	0	0	0	1
Radiographs	4	0	4	0	0	0
X-ray, CT, DXA and MRI	1	1	0	0	0	0
Any (not specified – focused on AI more broadly)	54	2	2	52	0	0
Patient pathways (including condition/s) (overall)						
Chest/thoracic/lung	19	1	4	2	14	0
Musculoskeletal	1	0	0	1	0	0
Bone maturity	1	1	0	0	0	0
Prostate	5	1	1	2	3	0
Colorectal	7	0	1	1	5	2
Breast	3	0	0	0	3	0
Coronary artery disease	3	0	0	0	3	0
COVID-19	1	0	0	0	1	0
Dentistry	2	0	0	1	1	1
Stroke	2	0	0	0	1	1
Pulmonary embolism	2	0	0	0	2	0
Varying: Pulmonary embolism, intercranial haemorrhage and acute cervical spine fractures	1	0	1	1	0	0
Thyroid	3	0	0	0	3	0
Skin	1	0	0	1	0	0
Fractures	4	0	2	0	3	0
Anaesthesia	1	0	0	0	1	0
ACL ligament ruptures	1	0	0	0	1	0
Varying: Cardiac, pulmonary and musculoskeletal	1	0	0	0	1	0
Lumbar spinal stenosis	1	0	0	0	1	0
Pneumonia	1	0	1	0	0	0
Possible gastric neoplasm	1	0	0	0	0	1

	Study characteristics	Overall (*n* = 114)	Study focus[a] Implementation (*n* = 6)	Experience (*n* = 12)	Perceptions (*n* = 59)	Effectiveness (quantitative) (*n* = 43)	Cost (*n* = 5)
	Varying: chest/lung nodules, COVID, fractures, scoliosis, prostate, neuro/dementia	1	1	0	0	0	0
	Any (not specified – focused on AI more broadly)	52	2	2	50	0	0
Participant group							
	Radiology staff		4	4	26	0	0
	Wider clinical staff		1	4	4	0	0
	Residents/students/trainees		0	2	8	0	0
	Patients		0	0	8	0	0
	Members of the public		0	0	2	0	0
	Radiology staff, wider clinical staff, stakeholders and patients		1	1	1	0	0
	Radiology and wider clinical staff		0	1	0	0	0
	Radiology staff and radiology students		0	0	6	0	0
	Radiologists and computer scientists/industry and IT staff		0	0	2	0	0
	Staff (radiology and wider) and members of the public		0	0	1	0	0
	Staff (radiology and wider), patients and members of the public (including some carers)		0	0	1	0	0
	Diagnostic images and readers		0	0	0	43	0
	Diagnostic images		0	0	0	0	5

a The total will not always add up to 114 as some papers had multiple areas of focus (as shown above), were conducted across countries, covered different imaging modalities and patient pathways and recruited multiple stakeholder groups.

Appendix 2

TABLE 9 Drivers for change and example quotes

	Reported drivers for change	Example quotes
System-wide/national factors driving AI implementation	• Growing acknowledgement of the potential use of AI innovations • Political interest in AI • The political interest mainly focused on using AI to help ease NHS pressures (e.g. workforce shortages and significant waiting times)	• *'We see it cited more and more in the direction of travel for NHS England that they want to harness digital innovations and especially AI'* (Diagnostic programme manager) • *'The government are using it as a big drive at the moment as well to try and help, because obviously the NHS has been having a bit of a battering over recent years with a whole host of things because demands on it just outweigh what they can do. So, they are obviously trying to find ways to help'* (Consultant radiologist and trust lead)
Local factors driving AI implementation	• Impact of local needs and service priorities • Impact of existing local capacity • Previous experience of AI projects • *All of these factors:* Highlight that trusts have autonomy in driving change locally	• *'The [region] of England is one of the worst areas for lung cancer. Our statistics are kind of above all the national averages for incidents and first stage of diagnosis. We've got varied waiting times for chest x-ray reports across the region, mostly good but sometimes these can get behind'* (Consultant radiologist) • *'We've got huge workload pressures within imaging. We've got the community diagnostic centres opening, we haven't got enough radiographers to run the machines that they've got never mind the new machines that are coming in that need staffing. We've got issues with radiologists in terms of reporting. We rely on the outsourcing companies, you know, really heavily to help with our reporting across the [region]'* (Diagnostic programme manager) • *'So [place], and Dr [name], have agreed to take the lead on this, due to their experience, and they shared all the information the trusts needed to work out, you know, what funding – you know, how much they needed from a capital point of view, how much they needed from the revenue'* (AI project lead and managing director for network) • *'We have quarterly meetings with the network, and we informally discussed it at that network to see if the trusts would be interested. Because we don't speak for the trusts. We have to have them onboard. Each trust still has its own autonomy and can decide whether to take part in these projects or not. So, that's how we took it forward'* (Consultant radiologist)

DOI: 10.3310/GJAR2722

Health and Social Care Delivery Research 2026

Appendix 3

TABLE 10 Summary of barriers/facilitators influencing implementation, overall and across different stages of implementation, and example quotes

Stage of implementation	Factor influencing implementation	Example barriers	Example facilitators	Example quotes
Overall (through-out all time points and implementa-tion contexts)	Time	• Tight timeline • Challenges engaging relevant staff in short time • Time needed to understand environment and build engagement • Dependent on scale, complexity of the AI and staff involved		*'It's being able to make the time to do it on top of still doing your other work as well. So that's the challenging this is you don't give up your day job to be doing extra stuff. You find the time to do the extra stuff'* (Consultant radiologist and trust lead)
	Logistics involved in obtaining and processing governance approvals and tailoring these to the local context	• Variations across trusts – for example care pathways, availability of specialists (who reports scans) and (existing IT infrastructure) • Trust-level governance required • Integrating/testing tool (local systems varied) • Collaboration required between clinicians, suppliers, PACS/RIS		*'Probably the most frustrating thing about this safety sign-off is it goes through all of these committees and everything, and at the end of it there's no change to actually what we're doing. So, it's not actually added any value. There's a lot of documentation they write down and it doesn't seem to add any value as such, it's just a necessary hurdle'* (Consultant radiologist and network lead)
	Resources and capacity (including project management)	• Workload on top of day job • Limited use of dedicated project management • Lack of project management required non-specialist staff (including clinicians) to perform this role – removing them from rou-tine responsibilities and resulting in additional project management support being needed – potential of exacerbating case backlogs in the initial deployment phase, worsening operational and clinical outcomes in the short term. • Implications for delivery of clinical services • Support for collecting data • Additional integration costs	• Project management: dedicated project manager, • Project management: regular meetings (variably used)	*'It [having a project manager] would have made a massive difference because we are literally just treading water, and I feel we're making it up as we're going along. We kind of come together every Friday and like what's not been done. We don't know what needs to be done until someone comes to us and says, why have you not done this?'* (Reporting radiographer) *'What we've found, [trust name] who've done really well, they've got an actual project manager on the project. Which is why they are flying ahead and it's been really well co-ordinated. The other trusts have put PACS administrators or the radiology manager, as project lead and they're not project leads as such. So it's been difficult with the other trusts just to keep everything moving along'* (Network project manager) *'We do not have a project manager, we don't have one at all. So it's myself, I'm kind of doing the clinical side of things with the team on the shop floor. Between us we're trying to do that [manage the project] but I don't really have anybody I can lean on or delegate anything to. I'm just kind of plodding along making mistakes as I go along and trying to learn from them'* (Reporting radiographer)

continued

DOI: 10.3310/GJAR2722

TABLE 10 Summary of barriers/facilitators influencing implementation, overall and across different stages of implementation, and example quotes (*continued*)

Stage of implementation	Factor influencing implementation	Example barriers	Example facilitators	Example quotes
	Service expertise and championing (including previous expertise and experience)		• Local champions drove local processes amidst competing priorities • Local championing within trusts was seen as important to ensure suitable engagement across a diverse set of professional groups. • Expertise and experience helpful in facilitating implementation	*'We only allow trusts to, essentially, take part if they could identify a local champion. And by "local champion" we basically mean someone who is not me or [name], ideally a consultant who could be in charge at the local hospital of this software – so how they want to implement it in the hospital. They are responsible for getting it through any of their own local governance processes, and they are responsible for speaking with thoracic surgeons and getting that up and running'* (Network programme manager) *'So, I think that was a very big positive, because we are experienced and help with upskilling the other NHS trusts'* (Consultant radiologist) *'Well having done previous procurements for PACS and various things before I was very familiar with the process and in [their local area]'* (Consultant radiologist and network lead)
	Collaboration – Sharing knowledge and learning – within and between networks, trusts and regions	• Degree to which networks were established locally – for example how well network teams knew staff within trusts	• For example NHSE: 'honest broker – framework, facilitating network engagement and collaboration, 'trouble-shooting' • For example Network: engaged/identified relevant groups, supported planning, project management, data, evaluation, network-level leadership (e.g. leading procurement, and supporting tasks around early deployment) (especially if trusts had limited project management capacity) • Trust level: Significant commitment from Trust teams (e.g. clinicians, PACS and RIS managers, and service managers) and suppliers (who engaged actively with project meetings, document writing, governance processes, and testing and adaptation of the AI systems). • Responsive suppliers: key to planning, participating in meetings, developing documents, training • Networks sharing learning, knowledge and documentation to support progress – sense of 'community of practice'	*'I know [name] and I have had separate meetings outside of those meetings with the people from some of the other networks just to touch base on a few things that we're finding similar in terms of those legal challenges and things'* (Project manager)

DOI: 10.3310/GJAR2722

Stage of implementation	Factor influencing implementation	Example barriers	Example facilitators	Example quotes
	Monitoring and evaluating the implementation of AI	• Lack of systems to access service-level data • Insufficient resources, local capacity and time to populate data for benefits register • Lack of statistical expertise to analyse and interpret data • Lack of timely information governance agreements – limiting data linkage and collection • Data availability ◦ Challenges quantitatively assessing inequalities due to insufficient resources, varying opinions on availability of patient information stored in RIS system. Lack of certainty regarding how to evaluate performance in coexisting patient conditions for example COPD, obesity. Data on failure rates (AI unable to interpret images) but yet to analyse it ◦ Variability of local systems and uncertainty about where relevant data were held ◦ Incomplete data returns (e.g. absence of referral source and cancer diagnoses particularly) due to lack of centralised data systems • Challenges in data access ◦ For example lack of pathway indicator within patient records – not possible to link CT scan to correct referring CXR ◦ Challenges capturing data on turnaround time due to voice recording technology that does not record reporting time in RIS • Challenges in data linkage capability ◦ For example linking data across hospital sites and systems to capture complete patient pathway information. Different networks and trusts had varying levels of linkage capabilities. Separate PACs and RIS systems complicated this. Multiple PACs suppliers within network also complicated this. ◦ For example linking data across departments such as cancer services – data recorded in separate systems	• Having a plan for evaluation • Support with data collection and monitoring (e.g. from individuals in dedicated data roles, or from Health Innovation networks) • Data linkage capability ◦ In some networks there were some hospital trusts that had combined PACs and RIS systems that can link image information to information specific to the radiology department such as reporting time. For these trusts, data collection was easier than for trusts that had separate PACs and RIS systems. • Centralised data systems and network-level data collection and support • Data availability ◦ Data on safety and diagnostic accuracy held by AI supplier and used for post-market surveillance – sometimes shared with deployment sites (e.g. portal to enable sites to access data)	*'It's actually quite a complicated data request I think that's come out from the national team and organisations don't typically have analysts in place to do this work permanently'* (Consultant radiologist and clinical director) *'So, we've got the NHS England benefits stipulations that we need to collect, and that's been quite overwhelming for the trusts. We wouldn't normally do any analysis in such great detail, and I think a lot of the data that's been suggested isn't readily available to the trusts, so they've had to go away and ask for different reports to be set up or manually manipulate data to be able to get that information out of the systems'* (Diagnostic programme manager) *'The benefits-based lining was, I think, unrealistic of NHS England. I think it was very unrealistic. A lot of it is not going to be helpful to the Trusts and they'll never want to measure it again because it's not going to offer anything. I think if they'd not just been given a spreadsheet to record data, and they'd just done a bit of a session with the trusts to get them to understand why it was important, why they needed this information and how it would benefit them, they would have had an easier ride. There's no secret that nobody's been happy about it. It's been painful for every single trust. There's not one that found it easy'* (Network project manager) *'A lot of the networks just passed that [data collection] straight onto the trusts and said please provide this data back to us. The protocol for [site name] is we want to do this at a network level to raise our network [data] maturity and they'll be questions we should answer at network level. Realistically, you know, the trust wouldn't be able to do this so I supplied the data for the trust, they were happy with that'* (Consultant radiologist and network lead)

continued

TABLE 10 Summary of barriers/facilitators influencing implementation, overall and across different stages of implementation, and example quotes (*continued*)

Stage of implementation	Factor influencing implementation	Example barriers	Example facilitators	Example quotes
Procurement	Supplier factors	• Challenges from unsuccessful suppliers • Large number of documents submitted – hard to navigate • For suppliers – completion of extensive documentation not felt to be relevant to AI in health care		*'The lengthiest bit's been all the contesting legal challenges that came afterwards before everything got signed off, but the actual initial procurement aspect was quick, and it got signed off legally two, three weeks ago'* (Consultant radiologist and trust lead)
	Procurement scoring	• Openness of procurement questions • Difficulties agreeing key functions of AI within network • Submissions presented during procurement vs. actual implementation of AI software on local systems • Views that issues of Equality important but unlikely to change outcome of scoring	• Having a balance of different expertise on panel • Procurement expertise and experience	*'Because there were lots of questions asked that they just wouldn't know without – I can visualise what's happening. So, we were quite – we thought we were being very specific, but even then we learnt a lot, so we know how to be even more specific'* (Consultant radiologist)
	Regional and local navigation of processes	• Novel and complex process of implementing AI in local trust levels • Making decisions between network and trust levels	• Established networks better positioned to conduct engagement between networks and local teams, and had more knowledge of local specialists to recruit to procurement panel • Network teams supporting local procurement panels	*'The whole complexity of the procurement, the installation and the clinical safety, I think is a risk. As we know, there's so many companies offering this kind of product. There's going to be even more coming in the future. They are all providing very similar outcomes to their algorithm. It's going to be extremely difficult to pick which one's best and which one to go with'* (Consultant radiologist) *'I think you know, when you're talking about doing things at a network level it would be ideal that just one Trust would do that sort of work and then it's sort of you know, used by each Trust rather than everyone doing things independently. So, we've tried that but in reality I think each Trust is still you know, there are some things you can do as a network you know, to like you know, shared learning and things like that. But actually, you know, each Trust will want to follow their own procedures'* (Consultant radiologist)
	National support		• National team supporting local/regional teams with development of specifications (specification template, FAQ document, online workshop), and monitored procurement through weekly meetings	*'Having a framework, not feeling like you're doing it alone because we were part of an AIDF group with really good leadership in terms of being a very honest broker, so, I felt supported in that way'* (Lead of AIDF bid)
	Time	• Time constraints – arranging consensus meetings so everyone could attend (sometimes more than one meeting)		*'The procurement process that we had had a surprisingly tight timeline that was, sort of, over the Christmas period going into the new year, where there were also junior doctor strikes going on'* (Consultant radiologist) *'There were time constraints. It was all felt like it was rushed because they had to get this in for winter pressures'* (Project manager)
	Resources and capacity		• Procurement leads at network and trust level	*'I'm lucky that we have [name] who's our procurement lead. She was really helpful in guiding me so as what to say'* (Network programme manager)

DOI: 10.3310/GJAR2722

DOI: 10.3310/GJAR2722

Stage of implementation	Factor influencing implementation	Example barriers	Example facilitators	Example quotes
	Collaboration – Sharing knowledge and learning – within and between networks, trusts and regions		• Support from expertise within the wider organisation • Support from other networks (e.g. sharing documents)	*'You need to have a non-clinical team that knows what they're doing. Because I wouldn't have personally been able to do this all myself. It's not my area of expertise. I don't understand NHS procurement law and NHS management, so to speak. Yes, so you need to have a team that does understand it'* (Network programme manager)
Preparation for deployment	Time	• See above (overall barriers) ◦ For example local CS processes important but required time to navigate due to need to develop documentation (hazard logs), arrange processes for approval		
	Logistics involved in obtaining and processing governance approvals and tailoring these to the local context	• See above (overall barriers) • For example CS processes requiring adaptation for different trusts	• CS governance procedures and documentation for AI deployment required a high level of specialist resource.	*'You know, information governance, data sharing agreements, all sorts of that stuff and in our experience that takes a very long time. So, we're managing a lot of that'* (Consultant radiologist)
	Resources and capacity (including project management)	• Lack of dedicated trust-level project management – instead relying on member of (e.g. clinical or imaging) staff take on the project management role.	• Regular project meetings • Trust-level project management capacity (dedicated project managers – only available in a few trusts) • Network-level project management support to fill gaps in trust-level project management support	*'We have got the weekly one [meeting] which is every Monday. But we also hold one with the super users where we know they are busy so we might just pop them up a Teams chat asking would they like to call, have they got any issues, and I leave it to them how they prefer to feedback. Just so they are not waiting a week, if there are any issues they can contact us immediately really'* (Trust system implementation lead)
	Collaboration – Sharing knowledge and learning – within and between networks, trusts and regions	• Different IT systems across trusts	• Supplier engagement and IT team engagement to integrate AI with trust IT infrastructure • Standardised system at network level	*'They've [the supplier] been very good, I have to say, they've been really good at making themselves available. Obviously it's in their interest to make this work, but no, they've been a really good team to work with, actually'* (Trust lead and development manager for imaging division) *'The AI tool is going to fully integrate with the PACS and RIS suppliers, but we've got different suppliers, so it has been a challenge. We don't have one unified PACS and RIS across all sites. So we are working with all of the suppliers to integrate the AI tool so that the information flows to and from the existing software that we've got'.* (Diagnostic programme manager) *'[A colleague] said his experience of doing these projects is like skate-boarding on gravel and he said in [our network] it's like skateboarding on concrete. The concrete was actually that when we procured we had the foresight to procure PACS across the whole of the network, so we've got the same system and it's single instance, which makes it easier'* (Consultant radiologist and network lead)

continued

This article should be referenced as follows:
Ramsay AIG, Herbert K, Lawrence R, Sherlaw-Johnson C, Bagri S, Crellin N, *et al*. Implementing artificial intelligence in chest diagnostics for lung disease: A mixed-methods evaluation
[published online ahead of print March 18 2026]. *Health Soc Care Deliv Res* 2026. https://doi.org/10.3310/GJAR2722

DOI: 10.3310/GJAR2722

TABLE 10 Summary of barriers/facilitators influencing implementation, overall and across different stages of implementation, and example quotes (*continued*)

Stage of implementation	Factor influencing implementation	Example barriers	Example facilitators	Example quotes
Deployment	Logistics involved in obtaining and processing governance approvals and tailoring these to the local context	• Unclear governance • Teething issues with codes (if used for prioritisation) • Complexity of the settings into which AI was to be introduced	• Processes in place to ensure staff do not rely on AI	*'It opens up a whole new territory for medicine to deal with that they feel uneasy and inexperienced with and that is how to deal with discrepancy cause by AI analysis. And there's a big emphasis on who's responsible and really it's being treated in quite an unusual way because why would AI be different to any tool that we've used over the years? They are tools for the doctors to use, and of course they are responsible and it's just another tool in the armoury that they're using to give decision support. But I guess as AI progresses and becomes more autonomous, perhaps issuing reports, then there could be a shift in who holds full responsibility. So that's a big challenge, the responsibility and legality of it'* (Consultant radiologist and clinical director) *'We had issues [with the codes], so the reporting radiographers were happy to use the codes, they were happy with everything, the wording, everything, they were happy just to go along with it. We start using them and then the radiologists are coming like stop using them immediately we don't like the wording. It felt like that was, it was just stop, start, stop, start'* (Reporting radiographer)
	Resources and capacity (including project management)	• Challenges with capacity and NHS pressures • Informal 'superuser' members of staff as early adopters of AI platform proficiency, who then intended to roll training to other members of staff • Integrating AI into existing systems	• Staff having time to attend training • Capacity for staff to be 'superusers'	*'The main challenges will be trying to invoke hospital resources that are finite and that are not being prioritised for our project to carry out an implementation against a two-year contractually limited timeline and almost risking squandering the opportunity because you take too long in trying to deploy. That's the biggest risk'* (Network lead) *'I was briefly told that there was some sort of package on [platform] where you can access something, and it shows you how to use. I haven't found that yet and it's finding the time to do that, and that's next on my to do list but I haven't reached that far yet. So, I haven't had any training. I think that people need to have the allocated time to do it'* (Reporting radiographer) *'The AI tool is going to fully integrate with the PACS and RIS suppliers, but we've got different suppliers, so it has been a challenge. We don't have one unified PACS and RIS across all sites. So we are working with all of the suppliers to integrate the AI tool so that the information flows to and from the existing software that we've got'* (Diagnostic programme manager)
	Local staff characteristics		• Training of staff • Engagement and willingness from staff	*'It's just making sure they're not using the AI as a "oh, this is it. It's gospel". That's what, you know, what we're going to do' and base it on that. They should all, you know, hopefully, hopefully, have the sense to do that'* (PACS manager) *'There's almost a little bit of a thought that we're introducing this, I think it's going to have massive benefits, but people actually have to look, for it to be of benefit. So, you know, a wider sort of education around engagement with imaging'* (Trust lead and manager imaging division)

COPD, chronic obstructive pulmonary disease.

DOI: 10.3310/GJAR2722

Health and Social Care Delivery Research 2026

Appendix 4

TABLE 11 List of data metrics potentially useful for evaluation of AI for chest diagnostics

Outcome	Metric	Sub-category	Identified by	Comments
General volumes and workflow	Number of CXR/CT imaging tests performed	By source of referral (general practitioner, ED, etc.)	NHSE	Reason for referral may also be useful in some circumstances, although difficult to obtain.
	Number of CXR/CT imaging reports issued	By category of finding[a]	NHSE	
		By staff role (including outsourced providers)	NHSE (subdivision suggested study stakeholders and literature review)	
		By patient characteristic (relevant comorbidities and demographic characteristics)[b]	NHSE (subdivision suggested study stakeholders)	
	Number of CXR/CT images reported on by outsourced providers		NHSE	
	Number of CXR/CT images interpreted/assessed by AI (i.e. AI interprets an image and presents the outcome/findings)	By staff role	NHSE (subdivision suggested study stakeholders and literature review)	Determining which images make up this selection (if under 100%) may be difficult if not tagged for 'AI assistance in interpretation'.
		By category of finding[a]	NHSE (subdivision suggested study stakeholders)	
	Number of CXR/CT images that AI has failed to interpret	By category of finding[a]	NHSE (subdivision suggested study stakeholders)	Likely unrecorded by trusts for each image scanned, and maybe available from AI suppliers
	For CXR applications			
	Number of patients with a CT scan	By CXR finding[a] and patient characteristic (relevant comorbidities and demographic characteristics)[b]	NICE/NHSE	Linking patients' CXR results to CT scans may be difficult as there is no patient pathway indicator
	Number of CT scans for patients following a CXR reporting suspected LC	By patient characteristic (relevant comorbidities and demographic characteristics)[b]	NHSE (subdivision suggested study stakeholders)	
Impact on staff workflow efficiency and reporting/turnaround times	Staff time effort [in reporting on one exam (CXR/CT)]	By staff role	NICE	Unavailable in imaging data systems. Average reported in Benefit Metrics derived from department surveys

continued

TABLE 11 List of data metrics potentially useful for evaluation of AI for chest diagnostics (*continued*)

Outcome	Metric	Sub-category	Identified by	Comments
		By category of finding[a]	NICE (subdivision suggested study stakeholders)	
	Number of images (CXR/CT) auto-reported, i.e. AI-generated imaging reports are issued to the referrer, but retrospectively reviewed by clinicians within 24 hours		NHSE	Auto-reporting can skew analysis of report turnaround times, hence the need to count so they can be excluded from analysis
	For patients referred from ED: time between ED admission date/time and ED discharge date/time for patients with a CXR/CT	By reason for referral	NICE/NHSE	A lot of confounding factors determine delay in discharge from A&E, hence this metric may not reflect the isolated impact of imaging workflow
	Time between start of report and verification of report (CXR/CT)	By category of finding[a] and patient characteristic (relevant comorbidities and demographic characteristics)[b]	NHSE (subdivision suggested study stakeholders)	The time point for the start of the report is unrecorded in imaging data systems
		By staff role	NHSE (subdivision suggested by study stakeholders and literature review)	
	Time between CT scan date and CT report issued date	By category of finding[a]	NHSE	
		By patient characteristic (relevant comorbidities and demographic characteristics)[b]	NHSE (subdivision suggested by study stakeholders)	
	For CXR applications			
	Time between CXR referral date and CXR report issued date	By category of finding[a] and patient characteristic (relevant comorbidities and demographic characteristics)[b]	NHSE (subdivision suggested by study stakeholders)	
	Time (days) between CXR test date and CXR report issued date	By category of finding[a] and patient characteristic (relevant comorbidities and demographic characteristics,)[b] also whether auto-reported	NICE/NHSE (subdivision suggested by study stakeholders)	
	Of CXR reports which reported suspected LC, % reported within 24 hours of CXR scan		NHSE	
	Time (days) between CXR report and CT scan	By category of CXR finding[a]	NHSE	

DOI: 10.3310/GJAR2722

Outcome	Metric	Sub-category	Identified by	Comments
	Time (days) between CXR scan date and CT report issued date	By category of finding[a] and patient characteristic (relevant comorbidities and demographic characteristics)[b]	NICE	
	Proportion of CT scans issued to patients on the same day as CXR report issued	By category of finding[a]	Interviews and discussions with implementation sites	
	Time (days) between first CXR scan date and start of treatment for those with diagnosed cancer[c]		Suggested by study stakeholders	
	For CT applications			
	Time between chest CT scan and issue of report	By category of finding[a]	NHSE	
	Average follow-up CT scans per patient within 30 months of first scan which reported a lung nodule		NHSE	
	Time between first CT scan date and treatment start date for those with diagnosed cancer[c]		NICE	
Impact on diagnostics	Number of patients diagnosed with LC	By AI/human agreement[d]	NHSE	The numbers on diagnoses and stages are influenced by concurrent initiatives like the LC screening programme. It would be useful if changes could be linked to the influence of AI.
	Number of LCs by cancer stage	By AI/human agreement[d]	NHSE	
	Number of complaints/serious incidents associated with false negatives (patients with cancer who were undiagnosed)		Suggested by study stakeholders	Unclear whether NHS Resolutions will have these data disaggregated for LC patients
	AI/human reader agreement rate		NHSE	
	For CXR applications			
	False positives. For CXR reports issued in the period, % of cases which reported suspected LC reported as normal/no cancer in CT scan	By AI/human agreement[d]	NICE	

continued

This article should be referenced as follows:
Ramsay AIG, Herbert K, Lawrence R, Sherlaw-Johnson C, Bagri S, Crellin N, *et al.* Implementing artificial intelligence in chest diagnostics for lung disease: A mixed-methods evaluation
[published online ahead of print March 18 2026]. *Health Soc Care Deliv Res* 2026. https://doi.org/10.3310/GJAR2722

DOI: 10.3310/GJAR2722

TABLE 11 List of data metrics potentially useful for evaluation of AI for chest diagnostics (*continued*)

Outcome	Metric	Sub-category	Identified by	Comments
	False positives. For CXR reports issued in the period, % of cases which reported suspected LC not diagnosed with LC after biopsy[c]	By AI/human agreement[d]	Suggested by study stakeholders and literature review	
	False negatives. Proportion of patients with LC whose earliest CXR (in the previous 26 weeks) was not reported as suspected cancer		NICE/NHSE	
	Positive predictive value of CXR (for cancer diagnosis), that is for CXR reports issued in the period, % of cases which reported suspected LC with a positive cancer diagnosis (through biopsy)[c]		Literature review	
	Positive predictive value of CXR (for CT result), that is for CXR reports issued in the period, % of cases which reported suspected LC with a positive CT result		Literature review	
	For CT applications			
	False positives. Of CT reports issued in the period, % of cases which reported suspected LC not diagnosed with LC (through biopsy)[c]	By AI/human agreement[d]	Literature review	
	False negatives. Proportion of patients with LC whose earliest CT (in the previous 26 weeks) was not reported as suspected cancer	By AI/human agreement[d]	NICE/NHSE	
	False negatives- Proportion of patients re-referred for a CXR whose earliest CT (in the previous 26 weeks) was not reported as suspected cancer		Suggested by study stakeholders	
	Positive predictive value of CT (for cancer diagnosis), that is of CT reports issued in the period, % of cases which reported suspected LC with a cancer diagnosis (through biopsy)[c]		Literature review	

Health and Social Care Delivery Research 2026

Outcome	Metric	Sub-category	Identified by	Comments
Personal impact on patients and clinicians	Self-confidence and competence among staff	By staff role	Literature review	Collection of these data can be facilitated by the trust or drawn from studies/surveys with the appropriate sample
	Fatigue among staff		Literature review	
	Trust among patients		Interviews and discussions with implementation sites	
	Anxiety among patients	By category of finding[a]	Interviews and discussions with implementation sites	

a Suspected cancer/no cancer or prioritisation category. Categorisation depends on the purpose of AI and the process.
b May need linkage with other clinical records data to determine comorbidities.
c Will require linkage between administrative imaging data and cancer registry data.
d Maybe particularly useful where the AI identifies abnormalities that the human reader is not able to detect.

DOI: 10.3310/GJAR2722

Appendix 5

TABLE 12 Costs of AI implementation for chest diagnostics

Cost	Details of cost	Source of funding/cost-savings	Data availability	Other observations (if applicable)
AI tool	• Fixed setup and operating costs per CXR/CT scan, based on past annual trust/network activity • Additional software/hardware costs required for integration (beyond supplier agreement)	• Existing budgets – If IT support for AI integration/maintenance not provided through 'in-house-funds'	• Participating trusts • Due to commercial sensitivity, use restricted to aggregate values across diagnostic imaging type (e.g. CXR/CT scan)	
Staff resources/costs	• Potential staffing changes in number, type, time or responsibilities (e.g. shift from radiologists to radiographers; or project management responsibilities for clinicians)	• Staffing costs from existing departmental budgets. • No funding sources identified for additional deployment responsibilities	• Details of staff recruitment readily captured	• Any savings realised through changes in staffing may be initially mitigated by resource requirements for staff training or upskilling
	• Staff training/upskilling	• Training/upskilling activities performed as part of role/skills development. No additional funding resource allocated despite the need to release reporting clinicians from normal duties for training and platform familiarisation, prior to live use of AI	• Staff time/resource may be acquired anecdotally (e.g. via staff interview/survey)	• Some Networks have established informal 'superuser' staff as early adopters, who then roll out training to others, although time for adopting this role is not protected • Capacity for staff to accommodate this role in addition to existing duties varies across Trusts, and thus additional resource may be required to ensure efficient roll-out of training/upskilling and avoid adverse operational or clinical impact
	• Additional reporting capacity and costs (e.g. IT support, office space) may be needed for increased case throughput, especially with current case backlogs	• No additional resource requirements confirmed at time of study. • Likely sources would be departmental/IT/estates budgets	• Details of infrastructure/equipment expenditure and changing to staffing costs readily captured	
	• Possible 'report-from-home' system for after-hours and urgent reporting, with costs for IT infrastructure, secure workstation setup, and updates to staff contracts for hours and on-call responsibilities	• No resource requirements confirmed at time of study. • Departmental/IT budgets are the most likely sources of funding for implementation	• Details of infrastructure/equipment expenditure and changing to staffing costs readily captured	
	• Establishment of CS governance procedures and documentation for AI deployment (usually high-level of specialist resource)	• No AI deployment-specific resources identified. Governance procedures were completed through existing specialist corporate/clinical capacity	• Additional (clinical) governance capacity within imaging departments would be readily captured. General (trust- or network-level) governance capacity resource would be more difficult to extract from their overall corporate activities	

Cost	Details of cost	Source of funding/cost-savings	Data availability	Other observations (if applicable)
Non-staff resources/ costs	• Potential cost-savings from early diagnosis (incl. late-stage treatment, survival, recurrence, complexity, abnormalities, turnaround times)	• Some long-term savings in the reduction of representations for imaging from false-negative cases • Cost-savings likely to be realised outside of diagnostic imaging departments (e.g. treatment pathway and health outcomes)	• Availability would be challenging, as these metrics often occur outside diagnostic imaging settings and may be stored in unlinked data systems	
Operational resources/costs	• Direct AI deployment costs • Ongoing maintenance and subscription fees for cloud-based AI services • Data storage costs • Costs related to maintaining the computational infrastructure and facilitate the integration of AI tools • Technical support costs for troubleshooting AI tools and ensuring uptime	• Costs incorporated into contractual agreement • No deployment-specific funding identified for additional IT infrastructure costs required to enable AI tool integration. Where required, funding likely to be derived from existing IT budgets	• Where required, details of additional infrastructure expenditure, readily captured via IT budgets	• Costs and resource requirement risks exist, where lack of AI technology expertise (integration requirements/functionality relative to the local diagnostic pathway) may lead to mismatch of an AI platform to the required role • Risks of exacerbating case backlogs and/or worsening short-term operational and clinical outcomes through the assignment of non-specialist project managers (e.g. clinicians) for deployment via removing them from their routine responsibilities and inefficient and/or difficulties in managing deployment workstreams/tasks
	• Increased service productivity from deployment of AI in diagnostic imaging	• Trusts envisaged that productivity increases and costs savings may be realised via:• Automation of clinical tasks (e.g. abnormality measurement and progression analysis) • Improved caseload processing, via AI-based case prioritisation • Efficiencies and consistency in report generation and results dissemination • Improved referral of positive cases (e.g. 'same-day' referral to follow-up CT scan)	• Clinician productivity (e.g. active consultation/reporting time) is not collected as a defined metric, either by participating Trusts nor can these currently be collected via the AI platforms. This may however be derived anecdotally (e.g. via staff interview/survey), or from aggregate data (e.g. reports completed per hour) • Challenging for downstream benefits from productivity improvements in diagnostic imaging, as these metrics would be located in other settings and may be stored in unlinked data systems	
	• Service performance monitoring	• Service performance tools are being established at a number of Trusts. Where present, the resource typically resides within the role of trust/network Data Analysts, or with a dedicated AIDF/ACID Programme Manager	• Staff time/resourcing for tool development and maintenance may be acquired anecdotally (e.g. via staff interview/surveys)	
	• AI identification of incidental (non-LC) abnormalities	• AI may also bring the potential for additional service and health benefits, through increased identification of incidental (non-cancer) abnormalities/conditions	• Challenging to collate data on downstream outcomes and benefits external to the cancer patient pathway	• These potential benefits are beyond the scope of the RSET evaluations as currently defined but may contribute to the overall cost-effectiveness of AI in chest diagnostic imaging

Appendix 6

TABLE 13 Key influences of phase 1 evaluation on phase 2 design

Phase 1 workstream	Phase 2 activity
Rapid systematic review	Contents of interview topic guides Potential quantitative measures Potential cost-effectiveness measures
Qualitative analysis of procurement and early deployment	Identification of potential interviewees Identification of relevant activities to observe Contents of interview topic guides
Analysis of measuring impact	Potential data sources Options for analysis design
Cost-effectiveness	Potential data sources Options for analysis design