# Mixed-methods Evaluation of Artificial Intelligence in Chest Diagnostics (Phase 2)

**Study protocol (v1.1, 19th March 2025)**


***Principal Investigator:***

*Dr Angus I G Ramsay, University College London (UCL)*

***Research team:***

*Chris Sherlaw-Johnson, Nuffield Trust*

*Dr Kevin Herbert, University of Cambridge*

*Stuti Bagri, Nuffield Trust*

*Malina Bodea, Nuffield Trust*

*Dr Nadia Crellin, Nuffield Trust*

*Holly Elphinstone, UCL*

*Dr Amanda Halliday, Public contributor*

*Nina Hemmings, Nuffield Trust*

*Dr Rachel Lawrence, UCL*

*Pei Li Ng, UCL*

*Joanne Lloyd, Public contributor*

*Dr Efthalia Massou, University of Cambridge*

*Raj Mehta, Public contributor*

*Prof Stephen Morris, University of Cambridge*

*Dr Jenny Shand, UCL*

*Dr Holly Walton, UCL*

*Prof Naomi J Fulop, UCL*



**FUNDING STATEMENT**

This project was funded by the National Institute for Health and Care Research, Health and Social Care Delivery Research programme (Ref: NIHR156380). The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the HS&DR, NIHR, NHS or the Department of Health and Social Care.

# 1 SCIENTIFIC ABSTRACT


***Background***

Artificial Intelligence (AI) refers to advanced technology that can perform complex tasks linked with human intelligence. AI has been used to support radiology in several clinical settings, including lung cancer detection and diagnosis. Evidence suggests that AI can contribute to accurate diagnosis, reduce errors, and improve efficiency. In June 2023, NHS England announced the Artificial Intelligence Diagnostic Fund (AIDF), which is funding 12 imaging networks of NHS Trusts across England to implement AI for chest diagnostics in 2024. Phase 1 of the evaluation led by NIHR RSET provided insights into the early implementation of these tools and developed a framework for monitoring and evaluation of AI tools for chest diagnostics in practice. However, there is limited evidence on implementation and use of AI in real-world settings, including staff experiences, patient and carer experience, effectiveness, and costs.

***Aims and objectives***

This rapid mixed-methods evaluation aims to explore the implementation, experiences, impact, and cost of AI tools for chest diagnostics.

The evaluation will address the following research questions (RQ):

1. How is AI for chest diagnostics being implemented and used in practice in NHS services in England?
2. What are the experiences of staff involved in delivering care supported by AI tools in chest diagnostics?
3. What are the experiences of patients and carers who had chest x-rays or CT investigations that were analysed by staff (supported by AI tools)?
4. What is the impact of using AI for chest diagnostics?
5. What are the cost and resource implications of setting up and delivering AI for chest diagnostics?
6. What are the lessons for future implementation and evaluation of AI in diagnostics?

***Methods***

This will be a mixed-method evaluation of implementation, experiences, impact, and costs of AI for chest diagnostics in NHS services in England. The evaluation will be informed by the Major System Change Framework [1].

To answer RQs 1-3, we will conduct 3 in-depth trust-level case studies and up to 9 light touch trust-level case studies. For in-depth case studies, we will work within three Trusts based in separate NHS Imaging networks (networks of NHS trusts within a geographic footprint with shared leadership to deliver real-time imaging, shared reporting worklists, and the multidisciplinary workforce required to achieve these objectives). We will interview up to 32 staff members (10-11 per Trust) involved in delivering or supporting delivery of AI for chest diagnostics, and up to 18 patients (up to 6 per Trust) and carers who have received a scan supported by AI chest diagnostics. We will also observe up to 30 meetings and analyse key documents relevant to implementation and impact of these tools. For light touch case studies, we will recruit up to 9 Trusts based in imaging networks where in-depth case studies are not being conducted. In these case studies we will interview the Trust service lead and the imaging network lead, and again analyse relevant documentation. Qualitative data will be analysed rapidly using Rapid Assessment Procedures and inductive thematic analysis. We will then conduct in-depth analysis using deductive thematic analysis.

To answer RQ4, we will work with the same three participating trusts to obtain site data from imaging and patient administrative data systems. In addition, we will use data sources that are currently accessible or known to be accessible, to assess outcomes at the other sites and for comparators. We will also investigate the possibility of acquiring post-market surveillance data from the AI suppliers or the trusts themselves. We will develop mathematical models using data from the sites together with other available evidence to address outcomes that will be hard to measure from this data alone and to inform more generalisable findings.

To answer RQ5, we will develop an economic model to estimate key costs and resource use associated with AI platform deployment, and for patient flow through the chest diagnostic imaging pathway. The model will be populated by accessible data as outlined for RQ4, to inform parameters on patient flow through the chest diagnostic imaging section of the patient lung cancer care pathway. Where data are not available directly from the participating trusts, published trust-level or aggregate data (e.g., regional/national) will be used to calculate proxy estimates (e.g., on a per-patient basis) for inclusion in the model. Data on the levels and impact of AI deployment on organisational resource use and costs (e.g., staff type/numbers, equipment, IT infrastructure) will be collated via a Trust-level data collection questionnaire which will be distributed to participating trusts. Where relevant, unit costs will be obtained from published sources [2,3], and inflated where appropriate to 2023/24 GBP, using NHS Pay and Prices Indices [2].

To answer RQ6, we will integrate findings from RQ1-5, together with findings from two online workshops with national stakeholders and staff involved in implementing AI diagnostic tools at network or trust levels.

***Patient and Public Involvement and Engagement (PPIE)***

Patients and the public have been and will continue to be central to this study. Our team includes the RSET PPIE co-lead and 2 public contributors with an interest in chest diagnostics. Our public contributors attend team meetings and have supported planning and writing of this protocol (e.g., contributing to planning discussions, and commenting on our protocol, research questions/design, recruitment documents, and interview topic guides). They will support our analysis (e.g., helping to interpret findings) and any outputs we produce (e.g., writing papers and presentations). Additionally, we held a PPIE workshop with four members of the public with experiences of and interest in these services to inform the development of this protocol; all were supportive of our proposed design and approach to recruitment and data collection. All PPIE involvement activities will be compensated in line with INVOLVE payment guidance.

**Stakeholder engagement**

The study has an Independent Advisory Group, which includes a range of stakeholders with relevant expertise, including patients, carers, and/or representatives of relevant charities.

***Timelines for delivery***

This rapid study will be conducted between November 2024 and December 2025, with the following milestones:

- *November 2024 – January 2025* – Develop study and submit ethics
- *February-March 2025* – Obtain ethics approval
- *April-October 2025* – Data collection and analysis
- *April-December 2025* – Write up findings and formative feedback
- *December 2025 onwards* – Submit final report, publish peer reviewed papers and other outputs.

***Anticipated impact and dissemination***

This evaluation builds on substantial engagement with relevant stakeholders (including NHS England, NICE, and relevant Royal Colleges). It will address important gaps in the evidence base highlighted by the NICE evidence generation plan for AI in radiology (published September 2023) and will develop recommendations on how AI tools for chest diagnostics can be implemented in future. Therefore, we anticipate there is potential for significant impact on policy and service delivery. To achieve such impact, we have developed an active dissemination strategy.

Dissemination methods will be discussed and agreed with stakeholders. We propose to share regular updates at national and network level established weekly meetings (e.g., AIDF weekly network meetings), other meetings where staff from trusts are present, and via the NHS Futures platform. We will also share findings through academic and professional-focused journal articles and conferences. We will produce accessible summaries of our findings, which may include slide-sets, blogs, and animations.

# 2 PLAIN ENGLISH SUMMARY

***Why is this study needed?***

Artificial Intelligence (AI) describes computer systems that can be trained to help solve problems. People think AI may help the NHS by improving accuracy of diagnosis and reducing workload and costs. NHS England have funded the use of AI for chest scans (e.g. x-rays and CT scans) in 66 NHS trusts (hospital organisations). Findings from phase 1 of our evaluation offered insights into the early implementation of these tools and found ways to study the impact and cost of these tools. However, little is known about the actual impact of AI tools, how much they cost and what staff, patients and carers think of them.

***What do we aim to do?***

We aim to study the use of AI for chest scans in practice by looking at:

- How they are used in NHS services in England
- What staff, patients and carers think of them
- Their impact on service delivery and the wider system
- Their cost

We will then develop recommendations to improve these services in future and guide further evaluations.

***How will we do this?***

We will use many different methods in this evaluation.

In-depth studies: In 3 hospitals, we will speak with up to 32 staff involved in delivering services, and up to 18 patients and carers who have received care supported by these services. We will also observe meetings and look at relevant documents.

Light touch studies: in up to 9 hospitals, we will speak with the local service lead and the regional imaging network lead and look at relevant documents.

To measure costs and impact on service delivery and the wider system related to using AI, we will analyse data from the sites that are using the AI and from other sources, including published data and services that are not using AI. We will look at effects on patients, services, and the wider system where possible.

We will hold two online workshops to share early findings and recommendations with national (e.g. policy makers, and service developers) and local stakeholders (e.g. staff involved in using these services), for feedback. This feedback will help guide our final recommendations.

***How have patients and the public been involved?***

When planning the study, we worked closely with our team's three public and patient representatives. We also held a workshop with four members of the public, who were supportive of our study focus and plans for collecting data. We will continue to gather feedback from members of the public throughout the project.

***When do we aim to complete the study?***

We will complete our study by December 2025.

***How will we share what we learn?***

We will share findings regularly with people who take part and others who are interested in the study. We will publish findings in academic journals, a final report to our funder, and a report to NHS England. We will also work with staff, patients, and the public to produce other ways to share findings with patients and carers (e.g. blog, short articles).

***Why is this research important?***

There is little known about the impact and cost of AI tools for chest scans, and how they have been used and how staff and patients/carers experience the care received. Our findings will help to develop recommendations that can inform future use of AI for chest scans in practice.

# 3 Introduction

## 3.1 Background and rationale

In recent years, policy documentation[4-13] and research evidence[14,15] have highlighted the potential for Artificial Intelligence (advanced technology that can perform complex tasks associated with human intelligence[14,16,17]) to support and transform healthcare in areas such as radiology. Research has indicated potential benefits regarding a range of outcomes (e.g. detection accuracy, error reduction/prevention, efficiency, decision-making and reducing workforce burden)[14,18,19]. However, there is mixed evidence for some outcomes (e.g. diagnostic accuracy)[20].

In June 2023, NHS England announced the Artificial Intelligence Diagnostic Fund (AIDF), which has invested £21 million to accelerate the deployment and implementation of AI diagnostic tools[21]. The fund focuses on chest x-rays and chest CT scans to improve the diagnosis of lung cancer and other conditions[21] and potentially help to address the current unmet need for faster chest x-ray reporting[21]. In the longer-term, the NHS propose that using AI to assist with the early detection of lung cancer can impact and improve patient care, with potential to improve patient outcomes[21]. However, NICE guidance has outlined numerous evidence gaps that must be addressed regarding the use of AI for chest diagnostics, including gaps on time saving and resource use, adverse effects, performance in different patient groups, ease of use and impact[4].

In November 2024, our team completed a phase 1 of a rapid evaluation of early implementation of AI tools for chest diagnostics deployed as part of the AIDF. The evaluation comprised a systematic scoping review of studies of AI in radiology[22,23] and an empirical study in 10/12 networks and 6/66 trusts implementing AI tools for chest diagnostics[23,24]. Findings from the phase 1 evaluation highlighted potential benefits of AI for chest diagnostics and in radiology more broadly. However, findings indicated varied implementation of AI tools for chest diagnostics in practice (e.g. in terms of the aims and purpose, referral and eligibility, and workforce models), together with challenges relating to the time taken and processes required to implement these tools[23,24]. Many factors influenced implementation (e.g. time, requirements to adapt for individual trust level, resources/capacity, importance of individual champions and expertise, collaboration and early engagement and shared learning, and support with data collection and monitoring)[22-24].

In terms of monitoring and evaluation, we identified several important evidence gaps relating to patient/carer experience, and evidence on cost effectiveness; due to gaps in data and limitations of data infrastructure[22-24]. These limitations are also making it difficult, in some sites, to gain a complete understanding of impact.

These gaps include limited evidence in literature regarding cost-effectiveness (only five studies evaluated costs), which poses significant challenges for the development of robust health economic models. Inconsistent linkage between patient records for the chest diagnostic services and those for downstream elements of the lung cancer care pathway (confirmatory diagnostic testing, treatment and long-term patient survival), presents an evidence gap for informing an economic model. This is compounded by the fact that the collection of long-term data (e.g., patient outcomes post-diagnosis), would not be possible with the time and resources available for a rapid service evaluation project.

In terms of the currently available data, fixed set up and operational costs per diagnostic procedure can be accessed through participating trusts, albeit as aggregated values due to commercial sensitivity; metrics related to staffing changes, such as recruitment, time spent and responsibilities, are captured within departmental budgets; infrastructure and equipment expenditures, like updating IT infrastructure to integrate AI into report-from-home systems, can be identified through departmental or IT budgets; potential long-terms savings from reduced late-stage treatments and false negative cases can be identified through linkages with data collected outside the diagnostic

imaging settings; while productivity related data, such as reports completed per hour, can be retrieved through staff interviews or a questionnaire.

The implementation of AI for chest diagnostics is operating across 12 imaging networks of NHS trusts, with potential to change organisation and delivery of care both within and across these organisations, with the aim of improving service delivery and outcomes at regional level. Therefore, these programmes may be conceptualised as examples of 'major system change'; defined as: "a coordinated, systemwide change affecting multiple organisations and care providers, with the goal of significant improvements in the efficiency of healthcare delivery, the quality of patient care, and population-level patient outcomes" [[25], p422]. Fulop et al's[1] framework of Major System Change outlines that it is necessary to evaluate all stages of implementation, including: the decision to change, decision on which model to implement, the implementation approach used, the implementation outcomes and the intervention outcomes. This framework has been used to evaluate several national system changes within the English healthcare system (e.g. reconfiguration of stroke services[1,26], specialist cancer services[27], COVID-19 remote home monitoring[28,29], and prenatal exome sequencing[30]).

Whilst research has shown that AI diagnostic tools have the potential to support and improve the detection of lung cancer, little is known about how effective and cost-effective these tools are, or how staff, patients and carers experience them. For recommendations to be made regarding the implementation of AI diagnostic tools, these knowledge gaps need to be addressed.[7]

## 3.2 Aims and objectives

This mixed-methods evaluation of Artificial Intelligence tools for chest diagnostics aims to address previous research gaps by exploring the implementation of AI tools or chest diagnostics, the impact and costs of implementing these models, the experiences of patients, carers and staff.

## 3.3 Research questions

1. How is AI for chest diagnostics being implemented and used in practice in NHS services in England?
2. What are the experiences of staff involved in delivering care supported by AI tools in chest diagnostics?
3. What are the experiences of patients and carers who had chest x-rays or CT investigations that were analysed by staff (supported by AI tools)?
4. What is the impact of using AI for chest diagnostics on service delivery and the wider system?
5. What are the cost and resource implications of setting up and delivering AI for chest diagnostics?
6. What are the lessons for future implementation and evaluation of AI in diagnostics?

Table 1 presents a summary of sub-questions within each research question.

## 3.4 Research team

The NIHR RSET team (AIGR, NJF, CSJ, MB, NC, EM, KH, RL, SB, HW, SM, PLN, HE, JS) and Public Contributors (RM, JL and AH) will deliver the independent service evaluation. The team works closely with national stakeholders and local teams (including implementation leads and clinicians); it is overseen by an independent project advisory group including researchers, policy makers, the voluntary sector, and patient/carer representation (see Section 12.3).

**Table 1. Research questions and sub-research questions for this evaluation of AI tools for chest diagnostics**

| Workstream | Research question | Sub questions | |
|---|---|---|---|
| 1 | RQ1. How has AI for chest diagnostics been implemented and used in practice in England? | a. | How AI is being used – what are the key functions, where it is being used in the diagnostic care pathway? |
| | | b. | How are staff (clinicians, managers, administrators) involved in using AI? |
| | | c. | How did early implementation work, e.g. in terms of planning and facilitation? |
| | | d. | How are patients informed about use of AI, e.g. in terms of communication/consent? |
| | | e. | How is AI for chest diagnostics being governed, e.g. in terms of information and safety? |
| | | f. | How have relationships with/between services, networks, and suppliers influenced organisation and delivery of AI? |
| | | g. | Have there been any adaptations in the service model, associated services along the care pathway, or governance over time? |
| | | h. | To what extent was AI for chest diagnostics implemented, e.g. in terms of uptake, spread, and fidelity |
| | | i. | How did implementation approaches (e.g. leadership, planning, and facilitation) and service models influence implementation outcomes (e.g. uptake, spread, fidelity) |
| | | j. | What are the implications for equity, diversity, and inclusion? |
| | | k. | Have there been any unintended consequences of implementing AI? |
| | | l. | What are the implications for sustainability? |
| | | m. | Which factors have been influential for implementation, e.g. functions of AI, patient groups, organisational context, network leadership, national programme/policy? |
| | RQ2. What are the experiences of staff involved in delivering care supported by AI tools in chest diagnostics? | n. | What are staff experiences of using AI for chest diagnostics? |
| | | o. | What are the factors (barriers/facilitators) that influence delivery of AI tools for chest diagnostics? |
| | RQ3. What are the experiences of patients and carers who had chest x-rays or CT investigations that were analysed by staff (supported by AI tools)? | p. | How have patients found the care received as part of the diagnostic pathway (including the use of AI to support diagnostics)? |
| | | q. | Are patients and carers informed/made aware of use of AI? If so, how? |
| | | r. | How are results communicated to patients? |
| | | s. | What are the experiences of patients and carers with different demographic and clinical characteristics? |
| | | t. | Which factors influence patient/carer experience of receiving care supported by AI tools in chest diagnostics? (e.g. trust, perceptions of AI) |
| | | u. | What could be done to improve patient and carer experiences? |
| 2 | RQ4. What is the impact of AI for chest diagnostics? | v. | What is the impact on patients, service delivery and the wider system? (Includes a) process outcomes like patient waiting times, processing times, knock on effects on the overall pathway, ease of use, resourcing and b) clinical outcomes like patient outcomes, safety, and diagnostic accuracy). What levels of AI performance will have a notable impact on these outcomes? |
| | | w. | What are the implications for patients with different demographic and clinical characteristics, e.g. co-morbidities, high-risk groups, socioeconomic status? |
| | | x. | How do data quality and data completeness affect the evaluation of impact? |
| | | y. | Which factors are likely to be most influential on impact, e.g. function of AI, IT infrastructure, patient profile |
| | | z. | How are implementation approaches and service models likely to influence impact (e.g. service delivery, patient outcomes) |
| 3 | RQ5. What are the cost and resource implications of setting up and delivering AI for chest diagnostics? | aa. | What are the implications for staff time, skill mix, and support in implementation and use (including tool maintenance) |
| | | bb. | How does the implementation of AI for chest diagnostics impact workforce requirements and workload distribution, including changes in staff roles, training needs, and potential shifts in resource allocation across diagnostic and operational workflows? |
| | | cc. | What are the wider resource implications of changes, e.g. support at Trust and network levels? |
| | | dd. | What are the costs associated with fee structure – product cost, deployment services, training, length of license? |
| | | ee. | What are the implications for equity, diversity, and inclusion? |
| | | ff. | Are there any unintended consequences, e.g. for workload? |

| | | | |
|---|---|---|---|
| 4 | RQ6. What are the lessons for future implementation and evaluation of AI in diagnostics? | gg. | Is AI for chest diagnostics sustainable? Which factors might influence this? |
| | | hh. | How transferable are the lessons to other healthcare diagnostics settings? |
| | | ii. | How did services/networks use learning from their local evaluation processes? |
| | | jj. | How might local/national evaluation support learning more effectively in the future? |

# 4 Methods

## 5.1 Design and theoretical framework

This is a multi-site rapid study that combines qualitative, quantitative, and health economic methods. This evaluation (phase 2) was informed by the findings from the phase 1 evaluation of AI tools for chest diagnostics)[22-24], scoping conversations (e.g. with academics, clinicians, policy representations, professional bodies, third sector organisations and regulators) and previous research. The evaluation will take place over 12 months (January 2025-December 2025).

The implementation of AI tools for chest diagnostics may be seen as an example of Major System Change. Therefore, this study will be informed by the Major System Change Framework[1], which was designed to understand the processes, outcomes and sustainability of such changes, in addition to the relationships between different stages of major system change[1] (see Figure 1 for a summary of design and theoretical framework).

We will employ a two-level case study design, including 3 in-depth case study Trusts and up to 9 light touch case study Trusts. We anticipate that the combination of these will ensure our evaluation can contribute both depth and breadth in its lessons. Table 2 summarises the main activities that will be conducted in in-depth and light touch case studies over the course of our evaluation.

**Table 2. Summary of activity in in-depth and light touch case study services**

| | In-depth case studies | Light touch case studies |
|---|---|---|
| Number of Trusts | 3 | Up to 9 |
| Workstream 1 activity | Staff interviews (up to 11 per service)<br>Patient and carer interviews (up to 6 per Trust)<br>Non-participant observations (up to 10 per Trust)<br>Documentary analysis | Staff interviews (up to 2 per service)<br>Documentary analysis |
| Workstream 2 activity | Local data sets (e.g. Radiology Information System (RIS) and Picture Archiving and Communication System PACs) | |
| Workstream 3 activity | Local data sets (Workstream 2)<br>Comprehensive data<br>Cost/resource questionnaire<br>Relevant material raised in Workstream 1 interviews | Cost/resource questionnaire |

## 4.1 Workstreams

The evaluation comprises four workstreams, outlined below.

**Workstream 1. Implementation, staff and patient/carer experience of AI tools for chest diagnostics**
This workstream will be led by AIGR with contributions from other RSET team members.

This workstream aims to answer research questions 1-3. In summary, this workstream will focus on implementation of AI tools for chest diagnostics in NHS services, staff experience with using AI tools for chest diagnostics, patient/carer experiences of receiving care supported by AI tools for chest diagnostics, and factors influencing implementation and experiences (see Table 1 for details of Research questions and sub-questions covered).

***Design***

Qualitative design, comprising semi-structured interviews, meeting observations and documentary analysis. See Table 3 for a summary of primary data collected within this workstream and later workstreams.

Figure 1. How workstreams will contribute to addressing Major system change framework components

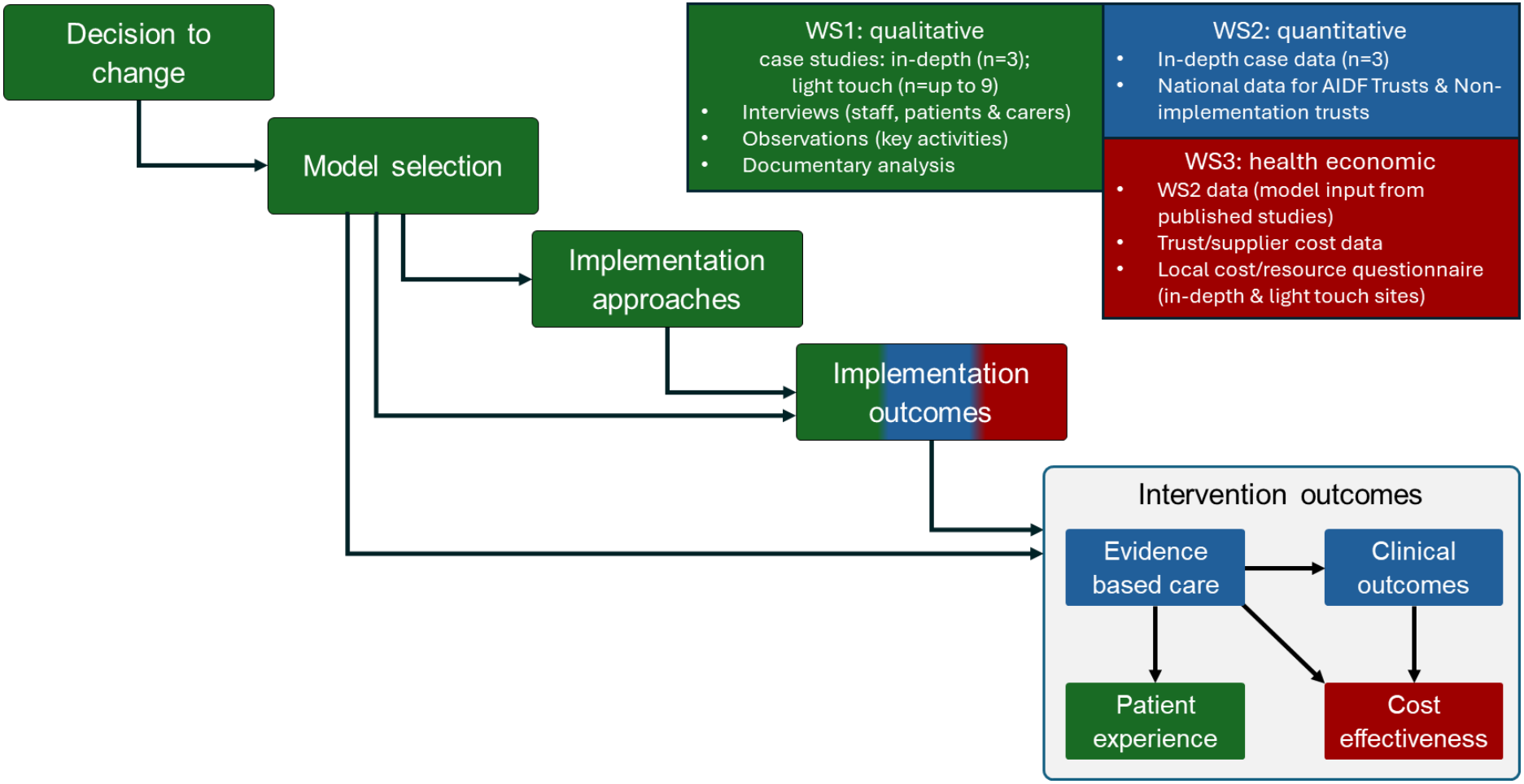


**Table 3. Summary of primary data collection methods within Workstreams 1 & 4**

| Workstream | Activity | Who we will invite to take part | Recruitment process | Who will conduct | Approx time | Recruitment study month |
|---|---|---|---|---|---|---|
| 1 | In-depth:<br><br>Interviews with staff | Up to 32 staff members (10-11 per trust). This will include:<br>• Staff with direct involvement in AI for chest diagnostics (e.g. radiologists, radiographers, suppliers, PACs/RIS mangers and suppliers)<br>• Members of the chest/lung MDT (e.g. doctors, nurses, oncologists, pathologists)<br>• Staff with wider oversight or experience of development (e.g. Information governance, or safety officers, data managers, project managers)<br>• Wider system staff (e.g. GPs or ED staff)<br>• Imaging Network lead | • Researchers work with trust leads to identify relevant staff/staff groups<br>• Researchers invite staff to participate via email<br>• Staff may also cascade study details to network<br>• Information sheet consent form sent, given time to review, participants consented, and interviews set up | RSET qualitative researchers (AIGR, NC, RL, NH) | 30-60 minutes | April-November 2025 |
| | Light touch:<br><br>Interviews with staff | Up to 18 staff (up to 2 per trust):<br>• Service lead<br>• Imaging Network lead | • As above | RSET qualitative researchers (AIGR, NC, RL, NH) | 30-60 minutes | April-November 2025 |
| | In-depth:<br><br>Interviews with patients/carers | Up to 18 patients and/or carers (6 per trust) who have received a chest x-ray or chest CT that was supported by the AI tools for chest diagnostics. | • Researchers work with service leads to identify patients/carers who meet eligibility criteria<br>• Trust staff share advert with eligible patients/carers<br>• Patients/carers asked to directly contact researchers to express interest, or ask trust staff to pass on contact details to researchers (who will follow up)<br>• Information sheet consent form sent, given time to review, participants consented, and interviews set up | RSET qualitative researchers (AIGR, NC, RL, NH) | 30-60 minutes | April-November 2025 |
| | Documentary analysis of trust-level documents | Relevant trust level documents pertaining to the implementation of AI for chest diagnostics (e.g. project plans, risk documents, meeting minutes, examples of anonymised AI reports, training materials, standard operating procedures, patient pathways, AI specifications) | N/A | RSET qualitative researchers (AIGR, NC, RL, NH) | N/A | April-November 2025 |
| | Observations of meetings | Up to 30 meetings relevant to implementation of AI tools for chest diagnostics. Meetings include:<br>• Project meetings,<br>• Training sessions<br>• Trust governance meetings, | • Researchers liaise with service leads to identify meetings to observe<br>• Information sheet and consent form sent to chair, given time to review, consent provided, and verbal consent checked with meeting participants at start of meeting | RSET qualitative researchers (AIGR, | Dependent on length of meeting | April-November 2025 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | • ICB oversight meetings | | | NC, RL, NH) | | |
| 4 | Online workshops | Two online workshops (n=up to 20 participants):<br>• Staff working in services that have implemented AI tools for chest diagnostics (n=1 workshop, 8-10 participants per workshop)<br>• Policymakers and other system leaders (n=1 workshop, 8-10 participants per workshop) | • Study adverts circulated via existing AIDF channels and networks, professional groups, social media, local third sector organisations, and direct invitation<br>• Interested individuals asked to contact researcher<br>• Information sheet consent form sent, given time to review, participants consented, and interviews set up | RSET qualitative researchers (AIGR, NC, RL, NH) | 60-90 minutes | October 2025 | |

***Sample***

*Site selection*

We will recruit 3 of the 66 trusts that are implementing AI to take part in this study as in-depth case studies. We will recruit up to 9 trusts as light touch case studies.

Selection of sites will be informed by our learning from the phase 1 evaluation[22-24] and an expression of interest process, whereby we will contact all trusts implementing AI for chest diagnostics (through the national meetings) and ask trusts to let us know if they would be interested in participating as a case study site for this evaluation.

We will purposively sample trusts as follows:

- All case studies (in-depth and light touch) will have implemented AI for chest diagnostics
- In-depth trusts will have good data availability (e.g. likelihood of obtaining local RIS/PACS data reports, and potentially linkage to cancer registry data) and good data quality to facilitate evaluation of impact and cost in workstreams 2 and 3
- Light touch trusts will be located across the 9 imaging networks in which in-depth cases are not being conducted.
- Within our case study trusts, we will seek to ensure representation across a range of characteristics of services and their contexts, including: the purpose of the AI tool (prioritisation, identification of lung cancer vs identification of other chest conditions), type of scan (chest x-ray vs chest CT), supplier and geographical location (e.g. urban, rural, coastal). Other relevant characteristics that will support decision making include: referral pathways, leadership approach (imaging network vs trust) and local PACs/RIS set-up (e.g. local arrangement vs regional platform).
- Sites that may meet the implementation and data quality criteria have been identified through engagement with local networks and by tracking implementation progress. Selection of specific sites will be guided by further discussion with imaging network leads and service leads in trusts identified as meeting these criteria.
- It should be noted that these sites will be 'early adopters' of AI technology in the context of chest diagnostics. This may represent a risk to developing lessons that may be translated to other settings. To help address this, we will seek to recruit some late-adopting services as part of our light touch sample. In addition, we will explore in our analysis and in our stakeholder workshops the extent to which our sampling represents a limitation to the learning generated, which we will discuss in any resulting reports from this evaluation.

*Interviews*

To explore implementation and staff experience, in our in-depth trusts we will aim to interview up to 32 staff members (approximately 10-11 per trust). Staff members will be recruited across a wide range of roles, including those with direct involvement in AI for chest diagnostics (e.g. radiologists – specialist and general, diagnostic and reporting radiographers, AI suppliers, PACs and RIS managers and suppliers, and teams who have been outsourced to provide reporting capacity), and members of the chest and lung multidisciplinary team (e.g. doctors, specialist nurses, oncologists, pathologists), and those with wider oversight of the role or development of the role (e.g. information governance teams, clinical safety teams, data managers, project managers, digital/AI leads, radiology physicist), primary care and emergency department staff where appropriate.

In our in-depth trusts, to explore patient and carer experience, we will aim to interview up to 18 patients and/or carers (approximately 6 per trust) who have had a chest x-ray or CT scan in one of the three participating trusts, and for whom AI for chest diagnostics supported their care delivery. We will aim to purposively sample patients and carers across a range of characteristics, including health outcome following review of scan (and therefore care pathway), and factors relating to socio-demographic characteristics (e.g. gender, age, ethnicity, disability). While we anticipate that these characteristics may influence how interviewees experience diagnostics supported by AI (e.g. many

sociodemographic characteristics may influence AI performance, with underserved groups disadvantaged), the purpose of sampling patients in this way is to ensure a range of perspectives, rather than to analyse experience by these characteristics.

In our light touch trusts, we will aim to interview up to 18 staff (up to two per trust), covering the local service lead and the regional imaging network lead.

*Meeting observations*

In our in-depth trusts, we will aim to observe up to 30 relevant meetings (up to 10 per trust) relevant to use and governance of AI for chest diagnostics and healthcare affected by AI for chest diagnostics. Activities to be observed will include AI implementation project meetings, AI training sessions, multidisciplinary team (MDT) meetings, directorate- and trust-level safety and quality committees, and regional oversight meetings.

*Documentary analysis*

In both in-depth and light touch Trusts, we will analyse local documents pertaining to the implementation of AI for chest diagnostics. Trusts will be asked to provide relevant documents, including: project plans, risk documents, meeting minutes, examples of anonymised AI reports, training materials, standard operating procedures, patient pathways, AI specifications, local audits and evaluation plans.

See Section 7 for recruitment processes.

***Topic guides***

Topic guides have been developed iteratively, informed by Phase 1 findings[22-24], scoping conversations, the Major System Change Framework[1] and previous research[15,31-33] (see Appendix 1 for topic guides).

**Interviews with staff** will cover the following topics: the interviewees' role and professional background, their views on the reasons and drivers for implementing AI tools for chest diagnostics, aims/purpose/function of AI tools, how AI tools are intended to be used and being used in trusts, how care is supported by AI in their trust, their experience of using AI for chest diagnostics (including training, support etc), perceived impacts and examples of perceived impacts, governance, data monitoring and evaluation, resource use, impacts on (in)equality, unintended consequences of using AI, barriers and facilitators to implementation and delivery, key learnings, and future use.

**Interviews with patients and carers** will take place in two parts. The first part of the interview will cover the following topics**:** the care they experienced, investigations and outcomes received to date (e.g., scan received, process of receiving their report), information provision (e.g. whether and in which ways they were informed about use of AI), the experience of the care they have received (things they liked, things they disliked), timeliness of care, barriers and facilitators to experience. As it is possible that patients may not be aware of the role of AI within their care pathway, we will then provide patients and carers with a short vignette that will explain how AI is being used in their local trust. We will then ask patients and carers about their views on AI, their knowledge about AI and how it can be used in healthcare, their awareness of AI within their care (e.g. whether it has been communicated to them/information provided), whether they would like to find out more about AI from their providers and if so, how, thoughts on benefits and challenges of AI, perceptions of impact of AI, unintended consequences of using AI, and how AI should be used in the future.

All participants will be asked to provide socio-demographic information at the end of the interview (NB sharing this information will be on a voluntary basis). For staff, this will include job role, length of time in post. For patients and carers, this will include health outcome (following chest x-ray/CT – if known), co-morbidities, age, gender, ethnicity, disability, sexuality, employment status.

Finally, we will give all patient and carer interviewees a secure link so that they can add any further information they would like to share in written format. This information will be transferred securely via REDCap directly to the UCL Data Safe Haven for analysis by the research team.

***Data collection***

**For all staff interviews**, researchers (AIGR, NC, RL, NH) will circulate an advert to staff involved in implementing and delivering AI tools for chest diagnostics at the three selected case study sites to ask if they would be interested in participating in an interview. Staff may share details of the study with other staff members who they think may be eligible/interested. Interested individuals will be asked to contact the researchers to express interest in taking part. Researchers will also contact key participants via email to invite them to participate. Researchers will then provide information sheet and consent forms.

**For patient and carer interviews,** sites will be asked to circulate study adverts to patients and carers who have received care that has been supported by the AI tools for chest diagnostics. Once interested individuals have contacted or been contacted by the researchers, researchers (AIGR, NC, RL, NH) will ask some basic eligibility criteria (e.g. whether they are over 18 and have received a chest x-ray or CT in one of the participating trusts since the implementation of AI) and will provide information sheets and consent forms. We will offer the option for information sheets and consent forms to be translated, if needed. Prior to the interview, participants will also be asked to answer some questions about their socio-demographic characteristics (either verbally during the interview, or by completing a short socio-demographic survey prior to/following the interview); sharing this information will be voluntary.

**All interested interviewees (staff and patients/carers)** will be asked to provide electronic, written or audio-recorded verbal informed consent prior to taking part in the interview. Potential interviewees will be informed that taking part is voluntary. The interviews will take place either online, or over the telephone. Each interview will last between 30-60 minutes, will be semi-structured, audio-recorded on an encrypted Dictaphone (subject to consent), transcribed verbatim by a professional transcription service, anonymised and kept in compliance with the General Data Protection regulation (GDPR 2018) and Data Protection Act (2018). Interviews will be scheduled to take place during regular working hours, as staff are not being compensated for their study participation. Participants will be informed that they are free to withdraw up to two weeks after the date of their interview.

For meeting observations, we will ask our point of contact for each trust to provide details of relevant meetings for the research team to attend. Researchers (AIGR, NC, RL, NH) will obtain electronic or written consent from the chair of the meeting and will also check verbally with members at the start of each observed meeting if they are happy for the observation to take place. Researchers will take anonymised notes on topics relating to the research questions using a standardised note template.

For documentary analysis, researchers (AIGR, NC, RL, NH) will ask relevant contacts at each trust to send required documentation.

***Data analysis***

To analyse findings from this workstream, we will use a medium Q thematic analysis approach[34], combining inductive thematic analysis and the use of a coding framework[35].

Data collection and analysis will be carried out in parallel, using Rapid Assessment Procedure (RAP) sheets[36], guided by research questions and the Major System Change Framework[1]. Qualitative data will be analysed by named researchers in the Rapid Service Evaluation Team (AIGR, NC, RL, NH). Researchers conducting interviews will take real-time notes and will input these notes into Rapid Assessment Procedure sheets following each interview. The categories used in the RAP sheet will be based upon the interview topic guides. There will be flexibility to add categories during the research process.

Once notes have been added to the RAP sheets, researchers will use inductive thematic analysis[35] to inductively code these notes and develop initial themes and sub-themes. This rapid analysis will be used to share interim findings with key stakeholders throughout the study.

Following the rapid analysis, an in-depth analysis will be undertaken. Researchers will use the initial themes and sub-themes developed during the rapid analysis to develop a coding framework. This coding framework will be applied to interview transcripts and observation field notes. The coding will be used to develop the final themes and sub-themes.

We will aim to undertake cross-case comparisons across the case study sites and staff characteristics (e.g., to explore barriers/inequities relating to implementation, delivery and patient experience). Interpretation of findings and write up will be discussed and agreed with the study PPIE group, advisory group and wider team prior to finalising.

**Workstream 2. The impact of AI tools for chest diagnostics**

This workstream will be led by CSJ with contributions from other RSET team members.

This workstream aims to answer research question 4. In summary, this workstream will focus on evaluating the impact of AI tools for chest diagnostics (see Table 1 for details of Research questions and sub-questions covered).

This study will be investigating outcomes from only a small number of sites, each coming from different starting points and with varying modes of implementation. These will be sites which have been implementing the AI for a sufficient period of time to see impact and where there would be enough data. They would also need to be sites where we can be assured of timely access to the data. Although we seek to be able to evaluate whether observed changes in these sites are significant, it is hard to obtain generalisable findings from their data alone. Moreover, likely availability of data and timescales will make it hard to observe longer-term impacts on, for example, cancer stages at diagnosis. Therefore, we will enhance the analysis with mathematical models of the chest diagnostic pathway which will be informed by data from the AIDF sites alongside other available evidence from published studies or the grey literature.

***Design***

Quantitative analysis of data derived from Hospital Episode Statistics (HES)/Diagnostic Imaging Database (DID)/Benefit Register from all trusts that have implemented AI for chest diagnostics through AIDF (n=63). This includes all networks involved in the AIDF programme, excluding sites whose deployment is for musculoskeletal conditions. More detailed analysis of imaging data (mainly RIS/PACs) from the three in-depth trusts specified in Workstream 1. Mathematical modelling supported by site data and evidence from existing studies of the chest diagnostic pathway.

Qualitative data collection and analysis to understand views from staff and other stakeholders on the factors and implementation approaches likely to influence impact.

***Sample***

For the quantitative study, the in-depth site sample is equivalent to that of workstream 1. If data is difficult to extract or is of particularly poor quality for any of those sites, we will approach a light-touch site to request local RIS/PAC data reports. More in-depth data from individual sites will also be sought from these same sites.

All 63 AIDF sites will be included for analyses of data sourced from HES, DID and the Benefit Metrics

Trusts that have not implemented AI for chest diagnostics will be included as comparators. DID and HES data is available for all trusts, so non-AIDF sites can be picked out from those data sources.

Mathematical models will be informed by site data and existing evidence from other studies.

***Measures***

The measures we will investigate will focus on:

- Caseloads and workflow (e.g. referral volumes, results of tests by category).
- False positive and false negative results (see below).
- Patient flow and waiting times (e.g. time from CXR to CT scan, time to confirmed cancer diagnosis).
- Image processing and reporting (e.g. turnaround times).
- AI use and performance (e.g. agreement with clinician, AI failure rates).
- Implications for different types of patient.
- Influence on early detection, i.e., the stage at which cancer is diagnosed.
- Influence on other outcomes in non-cancer pathways.

False positives will include follow-up CT examinations with negative results (for CXR applications only) and subsequent cancer diagnoses that are negative. False negatives will include cancer cases that are missed by the diagnostic imaging supported by the AI. In all cases, it is the accuracy of the human reader with AI available to support their decisions that is being measured rather than the AI alone, or just in cases where we know the AI has influenced decision making, as this would be very difficult to measure

***Data collection***

We will seek empirical data from up to three of the sites selected in workstream 1 for in-depth work. In addition, we will use data sources that are currently accessible or known to be accessible, to assess outcomes at the other sites and for comparators. These data sources include:

- The Diagnostic Imaging Dataset (DID)
- Hospital Episode Statistics (HES)
- NHSE Benefits Registers

Site data we seek will be downloads of aggregated or summary data from local data repositories of RIS or PACS systems. For some sites this may also be linked within to cancer diagnostic data.

DID data will be requested from NHSE and will include both AIDF and non-AIDF sites.

The evaluation team already have access to HES and have permissions to use it within all RSET projects.

The team also have access to the NHSE Benefits Registers via NHS England.

We will also investigate the possibility of acquiring post-market surveillance data from the AI suppliers or trusts, depending on local arrangements.

To inform our mathematical models of the process, we will use existing evidence from studies of the chest diagnostic pathway. These studies do not need to have investigated the use of AI.

***Data analysis***

The proposed use of each source of data and their pros and cons are shown in Table 4.

**Table 4. Proposed data sources for workstreams 2 & 3**

| Data source | Proposed use | Advantages | Disadvantages |
|---|---|---|---|
| NHSE Benefits Registers | Measures relating to case volumes, the change in processing times, times to follow-up tests, resourcing and cancer diagnoses. Measured at baseline and 6 and 12 months post-deployment.<br>Used for metrics not covered by other sources.<br>Mainly descriptive analysis. | Readily available for all sites.<br>Data submission a requirement of each site in order to obtain funding.<br>Ranges across most of the relevant outcomes. | Historic baselines.<br>Processing times measured as averages.<br>Not all data can be provided by all sites.<br>Concerns about complete recording of new fields such as image prioritisation categories.<br>Lacking all the granularity we might need. |
| Data from local sites | More detailed data that underpins the metrics in the Benefits Registers.<br>Within site exploration of differences among patients with different characteristics.<br>Within site comparisons, e.g. between patients with normal and abnormal imaging results.<br>Exploration of specific issues that may arise from workstream 1 interviews.<br>Findings from local studies. | Can ask for more granularity.<br>Can obtain more information on processing times than averages.<br>Can analyse data by patient characteristic.<br>Potential to analyse specific non-cancerous conditions such as infections and pulmonary embolism. | Since this data informs the Benefits Registers, the issues of data completeness are the same.<br>Agreements need to be put in place with each site we select.<br>Aggregated data may restrict levels of granularity due to suppression of low numbers. |
| DID | Analysis of outcomes relating to processing times. Comparison with non-AIDF sites. | Can be obtained for all sites and non-AIDF sites for comparison. | Only available for a limited number of outcome measures. |

| | | | |
|---|---|---|---|
| | | Data completeness and quality assessments are also issued alongside publication. | Aggregated data.<br>Approximate 5-month time lag.<br>Lacking all the granularity we might need. |
| HES | This depends on the outcome of the feasibility study. Potential outcomes include times between diagnostic tests and treatments. Analysis of differences due to patient characteristics. | Easily available.<br>Patient-level longitudinal data.<br>Can analyse data by patient characteristic.<br>Can analyse data for non-AIDF sites for comparison.<br>Can isolate GP referrals. | Feasibility is uncertain. It may some of lack the detail we need.<br>Diagnostic information in outpatient records may be limited.<br>Unable to distinguish suspected lung cancers from other reasons for CT referral.<br>Approximate 3-month time lag. |
| Post-market surveillance data (from suppliers or trusts depending on local arrangements) | Assessing performance of the AI tools such as clinician/tool agreement and AI failure rates. | Probably the only source of this data. | Agreements may need to be put in place with the suppliers. |
| AI supplier cost data | Key cost component for evaluation of AI deployment in WS3 | Cost data collated for all suppliers involved with the procurement process across all trusts in the AI deployment. | Aggregate estimates will not be fully representative of the specific costs applicable to the participating trusts |
| Published AI platform performance metrics (sensitivity and specificity) | Estimation of the rates of false positive and false negative results (where sensitivity and specificity data from sites, non-AI diagnostic imaging or AI suppliers are unavailable) | Published large sample study data | Published data may not reflect real-world service performance, given variations in care pathways and application of AI platforms between trusts |
| Participant trust questionnaires | Collection of resource use and cost data relevant to the AI deployment and patient diagnostic pathway (e.g., | Able to pose questions specific for the requirements of WS3 | Knowledge required for completing the questionnaire may |

| | | | |
|---|---|---|---|
| | staff type/numbers/time, equipment, IT infrastructure) | | not reside with one individual/group.<br><br>Some data will not have been prospectively recorded and therefore will be retrospective anecdotal evidence, raising the issues of accuracy, bias and generalisability |
| Published lung cancer patient pathway costs and outcomes, by stage of diagnosis or associated with false negative or false positive diagnosis | Will be used to populate sections of the care pathway which are beyond the scope/resource capacity of the RSET project team data collection plan. | Peer-reviewed estimates of costs and health outcomes for lung cancer patients in the UK population | The need to assume that the published data are generalisable to the patient sample in the participating trusts. |

The extent to which we can investigate these metrics will depend on what we are able to glean from these data sources, and their quality and completeness. Since many sites are in the early stages of deployment, this is not yet clear and the influence of data quality and completeness is included within the workstream.

For example, clinical outcome measures can only be obtained from sites where data is linked between radiology systems and cancer registries, allowing us to chart patients' diagnostic journeys to a definitive clinical outcome.

Understanding how outcomes differ for different types of patients, i.e., understanding implications of deployment on inequalities will require access to record level data available from HES. The value of HES in supporting this analysis will need to be explored first to understand its capabilities and limitations, so, to this end, we will undertake a short feasibility study supported by expert advice.

Where we plan to use DID or HES for comparators, we will use longitudinal data both before and after deployments and apply statistical methods that will account for deployment at different times. When using patient-level data, trust factors can be included as random effects and we could simultaneously explore any influence of patient characteristics. We cannot, however, account for simultaneous interventions that may be happening in comparator sites that aim to improve backlogs.

We will work alongside Workstream 3 to develop a model of the lung cancer patient pathway supported by site data alongside available evidence from published sources on diagnostic accuracy, including resource constraints and efficiency. These will map the progress of patients from initial tests through to any confirmed cancer diagnosis with progressions dependent on the underlying cancer stage. The purpose of these models will be to link different levels of AI performance to outcomes such as volumes of follow-up tests, missed cancer diagnoses and stage of cancer at first diagnosis. This would lead to more generalisable findings.

Once we scope data in HES and DID, we will know the degree to which we can investigate process outcomes (imaging processing, patient flow etc). After speaking with individual sites re their data linkage, we will know whether we can look at clinical outcomes and match radiology data with cancer registry. We will use existing evidence as well, but acquiring local data is important in centring this analysis around AIDF sites.

**Workstream 3. The cost and cost-effectiveness of AI tools for chest diagnostics**
This workstream will be led by KH with contributions from other RSET team members.

This workstream aims to answer research question 5. In summary, this workstream will focus on evaluating the cost and cost-effectiveness of AI tools for chest diagnostics (see Table 1 for details of Research questions and sub-questions covered).

The study will be investigating the resource use, costs and outcomes for deployment of AI in the diagnostic chest imaging stage of the lung cancer care pathway across a small sample of participating sites. This will be performed through the development of a pragmatically designed economic model, which will be populated by relevant data collated and synthesised from Workstream 2 (local datasets from the 3 in-depth participant trusts), responses to a participant questionnaire (completed by the 3 in-depth participant trusts plus up to a maximum of 9 light-touch participant trusts), relevant AIDF material (e.g., procurement documentation), and any relevant material raised in the course of the participant interviews (Workstream 1). Data required for populating the model which is otherwise unavailable from these sources (e.g., beyond the time or resource scope of this project), will be obtained from relevant healthcare datasets or published studies.

***Design***
A health economic evaluation, modelling resource use, costs and health outcomes for the deployment of AI within diagnostic chest imaging stage of lung cancer care pathway. The model will adopt a

decision tree approach with an NHS and Personal Social Services perspective, and a lifetime time horizon. Long-term costs and outcomes for lung cancer patients (by stage at diagnosis) will be derived from published estimates from relevant studies in a UK setting. Given the limitations of the resource capacity and timeframe of this project, will be designed pragmatically – being mapped to a simplified version of the National Optimal Lung Cancer Pathway (NOLCP). [Figure 2 [37]]

Patient flow and short-term costs and outcomes from the chest diagnostic imaging stage of the care pathway will be informed by data collected in this study, supplemented by evidence from existing studies for long-term costs and health outcomes.

The above-described model will be compared to the usual care pathways for chest diagnostics, which involve traditional diagnostic approaches without AI assistance. Data for these pathways will be drawn from the same sites pre-AIDF implementation, as well as from other non-AIDF sites where AI tools for chest diagnostics have not been implemented.

We will ensure that the comparator (usual care pathway) is as robust as possible. However, we acknowledge that the extent of the comparator will depend on the availability and quality of data from these sources. Evidence from existing studies will be sought to inform the parameters of the model where these data are not available in the study sites.

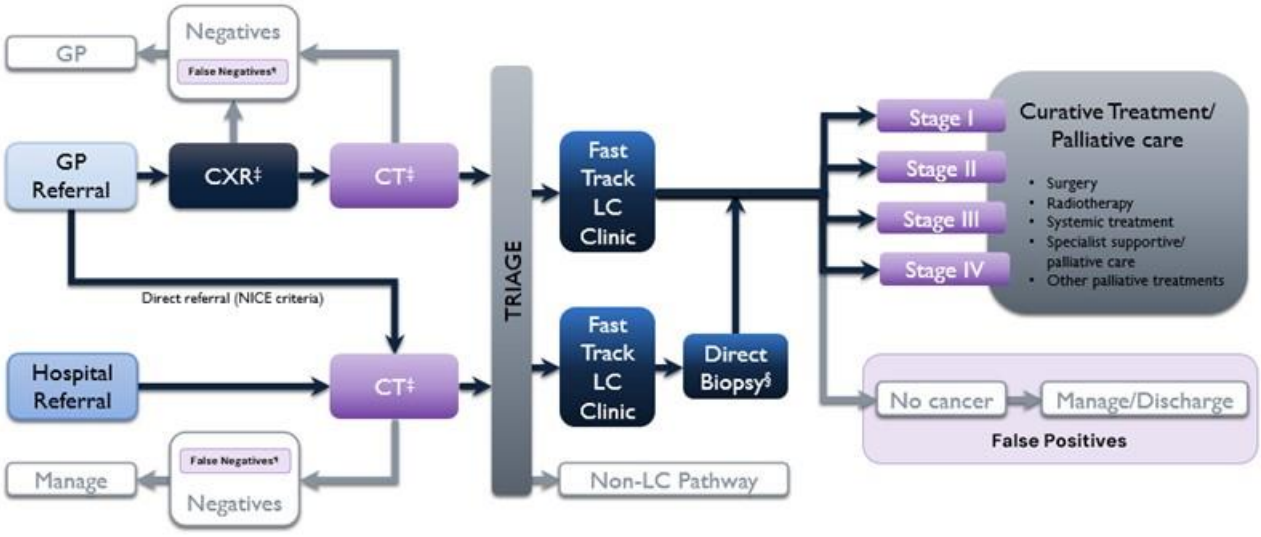
***Sample***

Further to the data collection based upon the sampling outlined in WS1 and 2, relevant data will be collated and synthesised from responses to a participant questionnaire (completed by both the 3 in-depth participant trusts plus up to a maximum of 9 light-touch participant trusts – site-specific data permitting).

Where AI/non-AI comparative estimates are used, these will be evaluated relative to baseline data from the respective sites, or (where available) from sites which do not use AI in the care pathway. The economic model will be further informed by existing evidence from other studies and for long-term (post-diagnostic imaging) costs and outcomes, and data from published studies and appropriate health service datasets.

Pre- and post-deployment data obtained during the collection period of the evaluation, will be converted to annualised estimates for costs and outcomes, weighted by the respective chest diagnostic imaging activity of the participating trusts.

**Figure 2. Overview of Lung Cancer Diagnostic Pathway.** †




† Adapted from National Optimal Lung Cancer Pathway, NHS England, 2020.
https://www.cancerresearchuk.org/sites/default/files/national_optimal_lung_pathway_aug_2017.pdf
‡Includes follow-up CT/CXR for patients with indeterminate results.

§ Rapid diagnosis pathway, where detailed staging and fitness investigations are not needed to guide management (*e.g.*, patients with advanced disease not suitable curative intent treatment).
¶ False negatives are presumed to re-present to either their GP or A&E.
Abbreviations: CT, computed tomography; CXR, chest x-ray; GP, general practice; LC, lung cancer.

### *Measures*

Relevant measures for the informing the economic model are outlined in Table 5. Those to be obtained (where possible) from this study, will include:

- Patient caseloads and workflow (e.g., referral volumes, results of tests by category).
- False positive results (based upon sites data or published specificity data for non-AI diagnostic imaging and from AI suppliers, where the former are unavailable).
    - Costs and outcomes subsequent to false positive results will be modelled on the assumption that patients progress for diagnostic testing and MDT confirmation of no lung cancer, prior to discharge.
- False negative results (based upon sites data or published sensitivity data for non-AI diagnostic imaging and from AI suppliers, where the former are unavailable)
    - Costs and outcomes subsequent to false negative results will be modelled on the basis that patients present to either their GP or A&E and are referred for diagnostic imaging following a delay after receipt of the false negative result, with a corresponding progression in stage of lung cancer at diagnosis.
- Any costs and outcomes associated with patient flow and waiting times (e.g., training for rapid delivery of results to patients, time from CXR to CT scan, time to confirmed cancer diagnosis).

Measures for the economic modelling obtained from published studies or appropriate published datasets, will include:

- Distributions of cancer diagnosis, by stage (after true positive or false negative results).
- Post-chest imaging costs for confirmation of diagnosis (e.g., multidisciplinary team case review, follow-up diagnostic and staging tests, biopsy)
- Costs and health outcomes associated with lung cancer by stage of diagnosis, or with a false positive or false negative result.

Costs and outcomes associated with other (non-cancer) conditions which may diagnosed via chest imaging, will be out of scope for this evaluation.

**Table 5. Key inputs for economic model[†]**

| | | Category | | | |
|---|---|---|---|---|---|
| | **Description** | **Care pathway** | **Cost** | **Outcome** | **Source** |
| GP referrals | Total referrals, CXR and CT (#) | ¶ | £ | | a |
| | Referrals to CXR (#, or % of total) | ¶ | £ | | a |
| | Referrals to CT scan (#, or % of total) | ¶ | £ | | a |
| | Positive predictive value, CXR (%) | ¶ | | | a, b |
| | Positive predictive value, CT (%) | ¶ | | | a, b |
| Hospital referrals | Total referrals, CT (#) | ¶ | £ | | a |
| | Positive predictive value, CXR (%) | ¶ | | | a, b |
| | Positive predictive value, CT (%) | ¶ | | | a, b |
| Chest imaging | CXR costs (AI, non-AI) | ¶ | £ | | a, c |
| | CT costs (AI, non-AI) | ¶ | £ | | a, c |
| | CXR sensitivity/specificity (AI, non-AI) | ¶ | £ | | a, c |
| | CT sensitivity/specificity (AI, non-AI) | ¶ | £ | | a, c |
| | AI imaging failure rate (CXR, CT) | | £ | | a |
| Post-imaging | Confirmatory diagnostic testing, true positive | ¶ | £ | D | a, b |
| | Confirmatory diagnostic testing, false positive | ¶ | £ | D | a, b |

| | | | | | |
|---|---|---|---|---|---|
| Fast-track LC clinic | Clinic MDT, no biopsy (#, or % of total) | ¶ | £ | | b |
| | Clinic MDT, biopsy (#, or % of total) | ¶ | £ | D | b |
| Curative treatment/ Palliative care | Treatment/care, by stage of diagnosis, true positive (#, or % distribution) | | £ | LY, Q | b |
| False Negatives | Patient re-presentation (GP, A&E) | ¶ | £ | D | b |
| | Chest imaging, confirmatory diagnostic testing, fast-track clinic (±biopsy) | ¶ | £ | D | a, b |
| | Treatment/care, by stage of diagnosis (#, or % distribution) | ¶ | £ | LY, Q | b |
| Other metrics | | | | | |
| Clinical safety | AI incidents (DATIX reports) | - | - | - | a |

† Inputs applicable for care pathways with use of AI in either CXR or CT scans.
Symbols: #, number of cases; %, percentage; ¶, patient care pathway input.
Abbreviations: A&E, accident and emergency; AI, artificial intelligence; CT, computed tomography; CXR, chest x-ray; D, disutility; GP, general practice; LC, lung cancer; LY, life years; MDT, multidisciplinary team; Q, quality adjusted life years.
(a) RSET study data; (b) healthcare datasets or published studies; (c) supplier data.

***Data collection***

Further to data relevant to the economic evaluation which will be collected in WS2 (above), data will be sought from the following sources:

Site-specific data on deployment, and AI use and performance (e.g., staff type/mix, image assessment and reporting time, AI failure rates) will be sought via a data collection questionnaire which will be distributed by the project team to trusts during the project data collection stage.

AI supplier cost data is available in a protected area of the FutureNHS web portal. Owing to its commercial sensitivity, these data will be aggregated prior to use/publication.

***Data analysis***

The proposed use of each source of data and their pros and cons are shown in Table 4. To address concerns about the comparability of sites, we will use techniques such as statistical adjustment for trust-specific and patient-specific characteristics (e.g., age, comorbidities, socioeconomic status) to control for differences between sites and populations. This approach will ensure that any observed differences in outcomes between AI and non-AI sites are not confounded by these variables

**Workstream 4. Development of lessons to inform future implementation and evaluation of AI for chest diagnostics**

This workstream will be led by AIGR with contributions from other RSET team members.

This workstream aims to answer research question 6. In summary, this workstream will focus on integrating findings from workstreams 1-3 and developing recommendations to inform future implementation and evaluation (see Table 1 for details of Research questions and sub-questions covered, and Table 3 for primary data collection methods).

***Design***

This workstream will integrate findings from workstreams 1-3, together with primary data collected during qualitative workshops.

***Synthesis of workstreams***

Integration of the workstreams will take place throughout the evaluation to enable complementarity of the workstreams: for example, workstream 1 has ensured interview topic guides cover issues relevant to workstreams 2 and 3; workstreams 2 and 3 will seek to analyse quantitative and resource data that is relevant to themes emerging from workstream 1; and workstreams 2 and 3 will collaborate on modelling work.

Findings from workstreams 1-3 will be synthesised (see individual workstreams) and then triangulated across workstreams Integration and triangulation will be facilitated by a number of processes: the methods used in WS1-3 have been developed by the team in order to be complementary, e.g. WS1 interview topic guides explore issues addressed in WS2 and 3, including perceptions of impact, resource use, and explanatory factors. Further, regular cross workstream meetings will enable qualitative, quantitative and health economic researchers to discuss findings and interpretations.

Findings from WS1-3 will be integrated around our research questions and the Major System Change framework. These findings and developing recommendations will then be presented at two stakeholder workshops.

***Stakeholder workshops***

***Sample***

We will hold two online workshops with up to 20 participants (8-10 per workshop). One workshop will be held with staff who have worked in networks and trusts that have implemented AI tools for chest diagnostics – including services that have done this outside the AIDF programme. The second workshop will be held with national stakeholders (e.g. commissioners, policymakers, system leaders, third sector organisations) with relevant expertise.

See Section 7 for recruitment processes.

***Topic guide***

During the workshop, we will first present a summary of findings developed from the phase 2 evaluation (workstream 1-3), together with an initial draft of lessons learned and recommendations for a) implementation of AI for diagnostics and b) evaluation and research to be conducted on such implementation. The workshop participants will then be involved in discussions relating to the following topics: i) their views on the findings, ii) their views on the lessons learned and any additional recommendations that should be added, and iii) their views on the future and sustainability of AI tools in a) chest diagnostics and b) radiology diagnostics more broadly. For some of the findings (e.g. the quantitative evaluation guide), we will ask stakeholders if they would be happy to comment on early drafts offline.

***Data collection***

Interested participants will be sent an information sheet and consent form in advance of the online workshops and asked to provide electronic or written consent ahead of the workshop. The workshop will be conducted by qualitative RSET researchers (AIGR, NC, RL, NH) and will take place online using Microsoft teams. Each workshop will last between 90-120 minutes. Researchers will audio-record workshop discussions on an encrypted Dictaphone (subject to consent) and take detailed notes to capture key findings. Recordings will be transcribed verbatim by a professional transcription service, anonymised and kept in compliance with GDPR 2018 and Data Protection Act (2018). Participants will be informed that whilst they can withdraw from the discussion, any data provided up until that point will be kept as it would not be possible to remove individual data from group discussions.

***Data analysis***

Following the workshop, researchers will analyse workshop findings using inductive thematic analysis[35]; organised around key themes and findings from workstreams 1-3. Workshop findings will support with the validation of, and further development of key recommendations for implementation and evaluation resulting from this work.

# 5 Study schedule

The planned timeline for the evaluation is as follows:

*Study design and develop protocol:* November 2024-January 2025

*Development of ethics materials and topic guides:* November 2024-January 2025

*Peer review of protocol:* December 2024-January 2025

*Protocol reviewed by NIHR:* December 2024-January 2025

*PPIE review of protocol:* December 2024-January 2025

*Publication of protocol:* March 2025

*Ethics approval:* February 2025-April 2025

*Data collection and analysis begins:* April 2025

*Data collection ends:* November 2025

*Data analysis ends:* December 2025

*Write up:* December 2025

*Submission of NIHR final report:* December 2025

*Summative dissemination (including summary slide set):* December 2025 onwards

The study Gantt chart is provided in Appendix 2.

# 6 Eligibility criteria

## 6.1 Inclusion criteria

**Staff interview participants**

- Local staff who work in or with the participating three trusts, and who are involved in organisation or delivery of care to patients receiving chest diagnostics which have been supported by the AI tools for chest diagnostics.
- Over the age of 18.
- English speaking or able to participate in an interview with an interpreter.
- Able to provide informed consent.

**Patients/carers**

- Patients and/or their carers (including family members) who have had a chest x-ray or CT scan that has been supported using AI for chest diagnostics, at one of the three trusts included in this study
- Over the age of 18.
- English speaking or able to participate in an interview with an interpreter.
- Able to provide informed consent.

**Workshop participants**

- National stakeholders with relevant job roles (e.g. policy makers, commissioners, system leaders, third sector organisations) relating to the implementation of AI, or local staff involved in implementing AI from the eleven networks and 60 trusts implementing AI for chest diagnostics as part of the AIDF.
- Over the age of 18
- English speaking or able to participate in an interview with an interpreter.
- Able to provide informed consent.

**Documentary analysis**

- Any documents pertaining to the implementation of AI for chest diagnostics at the participating three trusts.

**Meeting observations**

- Any meetings relevant to the implementation of AI chest diagnostics at the participating three trusts.

## 6.2 Exclusion criteria

- Anyone under the age of 18
- Anyone who cannot provide informed consent
- Patients/carers for which the AI tool was not involved in supporting their care

- Patients/carers at sites not included in this study

# 7 Recruitment and consent

## 7.1 Selection of case study trusts

To recruit participating trusts, we will present the study plans at existing AIDF network meetings and invite trusts to express interest in taking part. Sampling will be informed by findings from our phase 1 evaluation.[24] In addition, and to ensure a diverse sample, sites will be asked to provide some basic information to enable sites to be purposively sampled (e.g. data availability, the purpose of the AI tool, type of scan, supplier, geographical location, referral pathway, leadership approach and local PACS/RIS set up).

## 7.2 Initial identification

**Staff interviews**

The researcher (AIGR, NC, RL, NH) will work with leads at each site to identify potential staff groups at their trust that may be appropriate for interview. Researchers will contact potential participants via email to invite them to participate. Staff may also cascade details of the study (and an invite for anyone to contact the researchers if interested) to their staff networks to support recruitment.

**Patient/carer interviews,**

The researchers (AIGR, NC, RL, NH) will work with staff leads or R&D contacts at each trust. The staff leads (or research nurses, if available) will contact potential patients/carers who meet the eligibility criteria (either by telephone, email or post) to share a study advert and see if they would be interested in participating in the study. Potential participants will be asked to contact the research team directly if they are interested in participating; alternatively, potential interviewees may ask the staff lead/R&D contact to securely pass on their details to the researcher (using the secure UCL Data Safe Haven) if preferred. The researcher will then contact the patient/carer to provide further information.

In the first phase of our evaluation, the team learned that services are taking varied approaches to informing patients about the use of AI in the diagnostic process, with some sites choosing not to inform people explicitly. Therefore, the invitation to be interviewed may be the first time patients are made aware that AI supported their diagnostic process: this may cause patients concern or a desire for more information. To accommodate this eventuality, patients will be made aware of the purpose of this study and signposted to national and local sources of information at each stage of the identification/recruitment process, e.g. in invitation and recruitment documentation.

The team recognises that hospital services are extremely busy. Therefore, when recruiting in-depth sites we will ensure that the proposed approach to identification, invitation, and recruitment is feasible in these sites; further we will work with local research nurses in sites where they are available to support our work.

Note: for the purposes of patient and carer interviews, sites will be classified as Patient Identification Centres.

**Meeting observations**

The researcher (AIGR, NC, RL, NH) will liaise with staff leads at each trust to identify appropriate meetings to observe. For each meeting type, we will liaise with the lead of the event (e.g. meeting chair or lead trainer) regarding whether observation will be possible and appropriate.

**Workshops**

To recruit workshop participants, we will circulate study adverts via existing AIDF channels and networks, professional groups, social media, local third sector organisations, and direct invitation.

## 7.3 Informed consent (all data collection)

**All interviews and workshops:**

All potential interviewees upon expressing interest in participating will be sent a participant information sheet and consent form (either by email or post, depending on preference). Participants will be given at least 48 hours to review the information and consider whether to participate. If they are happy to take part, they will be asked to provide consent prior to the interview or workshop. An informed consent process using participant information sheets and written consent (scanned forms or typewritten/electronic signature), or audio-recorded verbal consent will be used for recruitment to ensure and demonstrate informed and voluntary participation. If participants would prefer to post consent forms back, they will be sent a pre-paid envelope to support this. If patients are not able/willing to take part in the interview but will still like their views to be included, we will ask patients if we can approach their carer (if they have one) to capture their perceptions of the patient's journey and overall experience with the service.

**For interviews**, the researcher will then arrange a time to conduct the interview over the phone or an online platform (Zoom or MS teams). If preferred or an individual interview is not possible, staff can choose to take part in a joint or group interview (where feasible). Similarly, patients and carers can choose whether they would like to take part in an interview separately or jointly.

**For the workshop,** participants will be sent the link to join once consent has been provided.

**Meeting observations**

We will send the chair or event lead the information sheet and consent form. If the event lead is happy in principle for the meeting to be observed, they will be asked to provide written consent, and information and consent forms will be shared with event attendees for information. Researchers will offer to present to the meeting an overview of the evaluation and what observations will involve. At the start of each meeting, we will also gain verbal consent from meeting/event attendees for the study team to observe and take anonymised notes.

# 8 PATIENT AND PUBLIC INVOLVEMENT AND ENGAGEMENT

We will continue to collaborate with our project PPIE members from Phase 1: Raj Mehta, Joanne Lloyd, and Amanda Halliday. They have been integral members of the study team since its inception, actively participating in the weekly project team meetings. Other PPIE engagement to date for Phase 2, has included: a virtual PPIE workshop with four members of the public (or family members of members of the public) who have lung conditions or experience with chest x-ray/CT scan and/or their family members. Additionally, we have a public contributor (HK) on our independent project advisory group. We have also engaged with representatives of several patient and public facing organisations and charities (Patients' Association, National Voices, Cancer Research UK, UK Lung Cancer Coalition, MacMillan Cancer Support, Patients Association, Ada Lovelace Institute, and Understanding Patient Data) when designing this evaluation. We have made several changes to the study protocol and design based on input from our PPIE and other engagement activity, including: the focus of our research questions, approaches to data collection, and contents of interview topic guides.

Throughout the study, public contributors on our project team (AH, JL, RM) will continue to attend project meetings, review study documents, and contribute to all aspects of the project, including interpretation of findings, co-authoring articles and accessible summaries, and other dissemination activities. We may also hold additional wider PPIE workshops to discuss findings and dissemination output more broadly.

Contributors are paid in line with NIHR INVOLVE guidance. We will regularly discuss our PPIE approach with public contributors to gain feedback on their experiences. PPIE activities will be supported by RSET PPIE leads (PLN, RM), Project Manager (HE), and project lead (AIGR) to ensure that all needs and preferences are considered.

# 9 EQUALITY, DIVERSITY AND INCLUSION (EDI)

To ensure that our project thoroughly and comprehensively considers issues of equality, diversity and inclusion, we will review compliance with our NIHR Rapid Service Evaluation Team (RSET) EDI assessment tool (see Appendix 2) at two stages during this project: (i) during development of the project evaluation and (ii) following data collection and analysis. The tool covers EDI considerations throughout the whole project, including when building the initial team, drawing on published EDI frameworks to consider EDI aspects relevant to the evaluation during the discovery and scoping phases, protocol development, stakeholder engagement, data collection, data analysis, and dissemination.

To date, the team have explored EDI considerations with our PPIE panel and with relevant stakeholders during scoping discussions, built EDI considerations into research question development, and have considered issues of EDI when developing this protocol (for example when considering site and participant selection).

# 10 FUNDING

The research costs for the study have been supported by the National Institute for Health and Care Research, Health and Social Care Delivery Research programme (RSET Project no. NIHR156380).

The study funding has been reviewed by the UCL Research Office and is sufficient to cover the requirements of the study.

# 11 DATA HANDLING AND MANAGEMENT

The study is compliant with the requirements of General Data Protection Regulation (2016/679) and the Data Protection Act (2018). All researchers and study site staff will comply with the requirements of the General Data Protection Regulation (2016/679) with regards to the collection, storage, processing and disclosure of personal information, and will uphold the Act's core principles. UCL, Nuffield Trust and University of Cambridge are joint data controllers and processors; the UCL Data Protection Officer is Alex Potts (a.potts@ucl.ac.uk). The data processors are AIGR, CSJ, ED, NC, EM, KH, SM, SB, RL, HW, NJF.

## 11.1 Data management

UCL will act as data controller for this study, with Principal Researcher Angus Ramsay leading on associated processes. He will process, store, and dispose of all data in accordance with all applicable legal and regulatory requirements, including GDPR and the Data Protection Act (2018) and any amendments thereto. Only relevant and necessary data will be collected, in line with the aims of this study. Data will not be transferred to any party not identified in this protocol and are not to be processed and/or transferred other than in accordance with the participants' consent.

In line with GDPR guidelines on data minimisation, we are only collecting personal data that is relevant and necessary for the purposes of this study.

**Qualitative data (workstreams 1&4)**
Participant interviews and workshops (qualitative data) will be recorded on an encrypted, password-protected digital recorder (only the researcher will know the password). Data will be collected by a team of qualitative researchers from RSET (AIGR, NC, RL, NH; University College London (UCL) and Nuffield Trust) (plus quantitative researchers from Nuffield Trust, and health economists from University of Cambridge, where appropriate topics regarding data and costs are discussed). Staff, patient and carer interview, observation and workshop consent forms, audio-recordings, anonymised notes and any documents received will be securely transferred using the Data Transfer portal onto the UCL Data Safe Haven (DSH, a secure electronic environment, certified to ISO27001 information

security standard and conforms to the NHS Information Governance Toolkit). Once transferred onto the UCL DSH, the data will be cleared from the Dictaphone.

Any participant consent forms received via post will be sent to our RSET team members at UCL and securely transferred onto the UCL DSH. Paper copies will be securely destroyed once scanned and uploaded to the UCL DSH. Electronic copies of consent forms received via email will be transferred onto the UCL DSH. If patients/carers would prefer to provide written consent or submit responses to socio-demographic questions using a secure survey link instead of verbally during the interview, the research team will develop online survey versions using the platform REDCap. Patients/carers would be sent a link and responses would be returned directly into the UCL Data Safe Haven via REDcap and would only be accessible by the qualitative research team (AIGR, RL, NC).

Digital audio-recordings of participant interviews and workshops will be sent to a UCL-approved contractor for transcription (TP Transcription Limited). Transcripts will be fully anonymised (names and places) and organised by participant codes. Anonymised transcripts and other relevant data will be stored in a secure folder to which only the named researchers (RSET qualitative team) have access. Only the research team will have access to participants' personal data (i.e., name and contact details). Participant identifier codes will be stored in the UCL DSH and kept separate from study data.

**Impact/cost data (workstream 2&3)**
Trust and/or Network-level benefits registers will be accessed via NHS England and transferred to the UCL DSH where it can only be accessed by members of the research team. This data is aggregated across the deployment sites and will, therefore, not contain any person-identifiable information.

Lung cancer diagnostics and outcome data from individual sites will not contain any person-identifiable information and will be processed within the UCL DSH (where it can only be accessed by members of the research team), after transfer from the network sites via the FutureNHS website**.** This data will either be aggregated with low numbers suppressed or in the form of summary statistics reflecting data distributions (e.g. medians, standard errors of turnaround times). Data Sharing Agreements will be drawn up for each site from which we request data.

Data from the Diagnostic Imaging Dataset (DID) will come from NHS England and in aggregated form. Again, this will be kept within the UCL DSH for analysis. An appropriate Data Sharing Agreements will be drawn up.

Hospital Episode Statistics (HES) outpatient, inpatient and emergency care datasets are held by the Nuffield Trust and stored on its secure server. An agreement is already in place with NHS England to allow Nuffield Trust staff to use this data for NIHR RSET projects. Staff undertaking patient interviews will have no access to this data and, similarly, all details on patient interviews will be stored on the UCL DSH in an area that will be inaccessible to staff using the HES datasets. This mitigates against interviewed patients being identified in HES.

AI supplier cost data will be held on DSH and will be accessed from a protected area of the FutureNHS. Data will be anonymised and aggregated to mean overall estimates for the respective diagnostic imaging method, prior to use. Returned participant questionnaires will be securely transferred onto the UCL DSH (electronic copies) and if applicable, paper copies will be securely destroyed once scanned and uploaded to the UCL DSH.

**Other (all workstreams)**

A Data Sharing and Processing Agreement is in place between the research team and NHS England for the purposes and duration of this evaluation. This covers documents held and/or developed by the AIDF programme.

# 12 PEER AND REGULATORY REVIEW

## 12.1 Peer review

This study protocol has been peer reviewed in accordance with UCL/UCLH requirements. It was peer reviewed by three reviewers external to UCL, with a diverse range of relevant clinical and academic expertise. It was also reviewed by the NIHR.

## 12.2 Ethics

Based on the Health Research Authority (HRA) decision tool and consultation with the UCL/UCLH Joint Research Office, most components of this evaluation (staff-focused qualitative work, quantitative work, and health economic work) can be classified as a service evaluation; we will submit these evaluation components for UCL ethical review. We will submit the patient and carer-focused qualitative work for HRA ethical review.

Although this is a relatively low-risk evaluation, we are aware of the sensitive nature of this work for organisations and individuals, especially patients and carers. The research team has experience in conducting health and care research on similarly sensitive topics. We will maintain the independence of the research, follow an informed consent process, and maintain the anonymity of participants and organisations.

## 12.3 Governance

This project is led by AIGR and delivered by a team of researchers and patient and public representatives. The research team meets on a weekly basis, with a set agenda that includes updates on progress of the AIDF programme, workstream-specific updates, project timeline, risk management, opportunities for dissemination and impact, and dedicated sections on PPIE and EDI. In addition, the project lead will report on progress to the RSET Executive Management Group monthly meetings, with a focus on progress, quality assurance, troubleshooting, and emerging learning and potential implications.

Independent oversight and advice will be provided in the following ways. Our project is supported by two independent clinical advisors. Secondly, a dedicated Evaluation Advisory Group, featuring independent stakeholders (including clinical, academic, and patient and public perspectives) will meet approximately three times at key stages of the study. Second, the study will be discussed at the RSET Stakeholder Advisory Board, which includes a range of clinician, academic, PPIE, and EDI experts, and meets every 6 months to offer oversight, challenge, and advice. Finally, we will update the AIDF evaluation subgroup regarding project progress and findings on a regular basis.

# 13 ASSESSMENT AND MANAGEMENT OF RISKS

## 13.1 Ethical considerations

During interviews and workshops, we will be asking participants to reflect on their views and experiences of AI diagnostic tools for chest diagnostics, and stakeholders may be hesitant to raise criticism or share information that is commercially sensitive. To address this, the participant information sheet will highlight that the research team are independent of those delivering the care service and that there are no right or wrong answers, and that the information will be fully anonymised (including names, places and particular AI tools). Our information sheet will also highlight escalation processes that researchers will follow should any safeguarding concerns arise during data collection. We will also emphasise that it is important to learn about the things that do not work as well, to improve these services for future patients. We will signpost participants to external services if required and where appropriate.

To minimise safety risks (e.g., spread of COVID) and reduce burden on participants (e.g., making/keeping arrangements for in-person interview), we will conduct interviews and observations remotely via MS Teams or Zoom.

## 13.2 Risks and mitigation

Potential risks and associated mitigations are highlighted in Table 6.

**Table 6. Potential risks and mitigation strategies**

| Workstream | Risk | Impact | Likelihood | Mitigation |
|---|---|---|---|---|
| All | Loss of key research staff | High | Low | There is a large project team. In the event of one member leaving there is capacity and resources for this person to be replaced from the wider team or to bring other researchers in. |
| All | Non-engagement from participating trusts | High | Medium | The research team has built relationships with networks and trusts throughout the course of the phase 1 evaluation. Additionally, we will continue to build these relationships throughout phase 2. We will put out an expression of interest so that sites taking part are keen to do so. Team members will have on-going meetings with site leads to discuss contribution required from each part during the evaluation. |
| All | Delays to ethical and governance approvals | High | High | We will ensure to submit ethical approval documentation in plenty of time to ensure that there is sufficient time for data collection. Additionally, we will work closely with participating trusts to ensure local approvals are expedited as quickly as possible. |
| 1 | Challenges recruiting patients, carers and staff for interviews | High | Medium | There is a risk the study may be delayed in recruiting participants because identification of patients / carers will be supported by trust staff. Similarly, staff would need to be willing/have time to participate in the evaluation.<br><br>We will discuss the practicalities and resource implications of the invitation/recruitment process with local sites up front when exploring the possibility of a service taking part in the evaluation (whether as an in-depth or light touch site). Activities will be specified in documents inviting sites to consider taking part, and will be explored further in meetings to confirm participation.<br><br>At each site, the team will identify a key point of contact regarding participation and will be in regular contact with them. The team will provide detailed information sheets to inform potential participants of the importance of the evaluation, why we have invited them to take part, their involvement and associated risks and benefits. |
| 2 & 3 | Limited availability of suitable data from sites | Medium | Medium | We will sample sites that have been implementing AI for a sufficient period of time to see impact and where there will be enough data. They also need to be sites where we can be assured of timely access to the data.<br><br>If we are unable to obtain adequate RIS/PACS data from sites, HES data may prove a sufficient back-up for some key metrics which is easily accessible, allows us to identify patient characteristics and, potentially, long term outcomes. However, this will depend on the outcome of our HES feasibility study.<br><br>We would also have to rely more on the NHS Benefits Registers with associated caveats around the strength of a comparator. |

| | | | | |
|---|---|---|---|---|
| | | | | If we also find that HES is insufficient then it affects our ability to look more deeply into the quantitative influence of different patient characteristics and inequalities on outcomes. We would then be only able to address these aspects qualitatively. |
| 3 | Small sample size potentially leading to identifiable commercial or patient data being reported in project outputs | High | Medium | Commercially sensitive metrics (e.g., AI platform costs) will be substituted for aggregated totals across all suppliers imaging type (CXR, CT scan).<br><br>Aside from HES, data for measuring effectiveness will be aggregated, with data suppression rules applied to avoid patient identification. Low number suppression rules will also be applied to any outputs of HES analysis. |
| 4 | Challenges recruiting participants or workshops | High | Low | The research team has built relationships with national stakeholders, networks and trusts throughout the course of the phase 1 evaluation. Therefore, we have sufficient networks to support the recruitment to these workshops. |

## 14 Recording and reporting of events/incidents

For this evaluation, we will complete a risk assessment that the research team will adhere to. Additionally, our evaluation will be registered with the UCL Data protection office.

Personal data breaches will be immediately reported to the UCL Information Security Group (ISG) and the UCL Data Protection Officer Alex Potts (a.potts@ucl.ac.uk) (as per form and guidance: https://www.ucl.ac.uk/legal-services/guidance/reporting-loss-personal-data). The following information will be provided: full details as to the nature of the breach, an indication as to the volume of material involved, and the sensitivity of the breach (and any timeframes that apply).

In the first instance, research participant complaints will be reported to the CI to investigate, as documented in the participant information sheet(s). Where appropriate, complaints will be submitted to UCL (via research-incidents@ucl.ac.uk, following the UCL Complaints from Research Subjects about UCL Sponsored Studies and Trials policy].

## 15 Monitoring and auditing

The project lead (AIGR) will ensure there are adequate quality and number of monitoring activities conducted by the study team. This will include adherence to the protocol, procedures for consenting and ensure adequate data quality.

Throughout the project, we will work closely with a range of stakeholders and also our project advisory group (see Section 12.3).

The research team will meet regularly throughout the duration of the evaluation. The evaluation will be discussed as a standing item at monthly NIHR RSET Executive Management Group meetings, in terms of progress against project milestones (see timeline and Gantt chart) and to address any practical or methodological issues.

To ensure that all researchers involved in data collection and analysis are supported throughout this project, we will build in time for reflection, debrief and discussions after data collection (i.e. workshops, observations, and interviews).

## 16 Training

The project lead (AIGR) together with the Director of NIHR RSET (Dr Jenny Shand) will review and provide assurances of the training and experience of all staff working on this study.

## 17 Insurance

UCL holds insurance against claims from participants for harm caused by their participation in this evaluation. Participants may be able to claim compensation if they can prove that UCL has been negligent. However, if this clinical study is being carried out in a hospital, the hospital continues to have a duty of care to the participant of the clinical study. UCL does not accept liability for any breach in the hospital's duty of care, or any negligence on the part of hospital employees. This applies whether the hospital is an NHS Trust or otherwise.

## 18 Archiving

The NIHR RSET team (UCL, Nuffield Trust, and University of Cambridge), and each participating site recognise that there is an obligation to archive study-related documents at the end of the study (as such end is defined within this protocol). The project lead (AIGR) confirms that he/she will archive the study master file at UCL for the period stipulated in the protocol and in line with all relevant legal and

statutory requirements. Study documents will be archived for a minimum of 5 years from the study end, and no longer than 20 years from the study end.

## 19 Publication and Dissemination

This evaluation builds on substantial engagement with relevant stakeholders (including NHS England, NICE, and relevant Royal Colleges). Our findings will address important gaps in the evidence base highlighted by the NICE evidence generation plan for AI in radiotherapy radiology (published September 2023) and will develop recommendations on how AI tools for chest diagnostics can be implemented and evaluated in future. Therefore, we anticipate this work has potential for significant impact on policy and service delivery related to AI tools for chest diagnostics and diagnostics more broadly. To achieve such impact, we have developed an active dissemination strategy.

Throughout the project, we will share findings (emerging and final) with key stakeholders (e.g. NHS England, NICE, royal colleges, relevant networks that have an interest in AI, sites implementing AI tools for chest diagnostics, and members of the public). To facilitate this, over the course of scoping and Phase 1, we developed a dissemination list covering over 100 individuals and organisations, with whom we have shared our findings to date.

To maximize the national and international impact of our evaluation, we will publish academic journal articles and present findings in a range of academic and profession-focused conferences. In addition, we will produce:

- slide sets covering key findings and lessons (e.g. see our Phase 1 slide set), to share with participants, stakeholders, and the wider public
- accessible summaries of our findings (e.g. blogs, explainers, and potentially short films or animations), to maximise public engagement with our findings
- bulletins promoted through our networks (e.g. the NIHR RSET newsletter, the Nuffield Trust Newsletter,
- outputs for newsletters and websites produced by the Royal College of Radiologists, the College of Radiographers, and AI and Digital Regulations Service
- We will work with relevant stakeholders to agree the content/format of these outputs

We will meet regularly with relevant stakeholders, providing updates at fortnightly meetings with the AIDF team and participating Imaging Network leads and monthly meetings of the National AIDF Service Evaluation Subgroup, and the National AIDF Strategic Oversight Board. We will also provide updates at trust/service-level meetings in sites where we are conducting evaluation work.

In addition, we will participate in events related to implementing AI in NHS services, including the National Diagnostic AI Forum, the AI Community of Practice, the NHS England AI Ambassadors Network, Responsible AI UK, and the NIHR-support incubator for AI and Digital Healthcare.

We will continue to build these networks throughout and beyond the lifespan of this evaluation in order to identify further opportunities to share our learning, and thus maximise the impact of this work.

## 20 REFERENCES


1. Fulop NJ, Ramsay AI, Perry C, et al. Explaining outcomes in major system change: a qualitative study of implementing centralised acute stroke services in two large metropolitan regions in England. Implementation Science 2015;11:1-13.
2. Jones KB, A. Unit costs of health and social care 2021. Personal Services Research Unit: University of Kent, Caterbury; 2022.
3. NHS England. National cost collection for the NHS: 2021/22: national cost collection data. 2022.
4. 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. (https://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).
5. National Institute for Health and Care Excellence. Artificial intelligence (AI)-derived software to help clinical decision making in stroke (https://www.nice.org.uk/guidance/dg57/resources/artificial-intelligence-aiderived-software-to-help-clinical-decision-making-in-stroke-pdf-1053876693445).
6. National Institute for Health and Care Excellence. AI technologies for detecting diabetic retinopathy. (https://www.nice.org.uk/advice/mib265/resources/ai-technologies-for-detecting-diabetic-retinopathy-pdf-2285965755558853).
7. National Institute for Health and Care Excellence. Evidence generation plan for artificial intelligence technologies to aid contouring for radiotherapy treatment planning. . (https://www.nice.org.uk/guidance/hte11/resources/evidence-generation-plan-for-artificial-intelligence-technologies-to-aid-contouring-for-radiotherapy-treatment-planning-pdf-17088176312773.).
8. NHS England. The NHS Long Term Plan. (https://www.england.nhs.uk/publication/the-nhs-long-term-plan/).
9. NHS England. Artificial intelligence (AI) and machine learning. (https://www.england.nhs.uk/long-read/artificial-intelligence-ai-and-machine-learning/).
10. Darzi A. Summary letter from Lord Darzi to the Secretary of State for Health and Social Care. (https://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).
11. Darzi A. Investigation of the National Health Service in England. (https://assets.publishing.service.gov.uk/media/66f42ae630536cb92748271f/Lord-Darzi-Independent-Investigation-of-the-National-Health-Service-in-England-Updated-25-September.pdf).
12. 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. (https://www.nhsconfed.org/publications/ai-healthcare#:~:text=The%20rapid%20expansion%20of%20artificial,overcome%20a%20range%20of%20challenges).
13. Labour Party. Wes Streeting speech at Labour Party Conference 2024. 2024.
14. Joy Mathew C, David AM, Joy Mathew CM. Artificial Intelligence and its future potential in lung cancer screening. Excli j 2020;19:1552-1562. (In eng). DOI: 10.17179/excli2020-3095.
15. 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(2):14604582211011215. (In English). DOI: 10.1177/14604582211011215.

16. Harwich EL, K. Thinking on its own: AI in the NHS. Reform Research Trust. (https://reform.uk/wp-content/uploads/2018/11/AI-in-Healthcare-report_WEB.pdf).
17. Istasy P, Lee WS, Iansavichene A, et al. The Impact of Artificial Intelligence on Health Equity in Oncology: Scoping Review. J Med Internet Res 2022;24(11):e39748. (In eng). DOI: 10.2196/39748.
18. Liu M, Wu J, Wang N, et al. The value of artificial intelligence in the diagnosis of lung cancer: A systematic review and meta-analysis. PLoS One 2023;18(3):e0273445. (In eng). DOI: 10.1371/journal.pone.0273445.
19. Chiu HY, Chao HS, Chen YM. Application of Artificial Intelligence in Lung Cancer. Cancers (Basel) 2022;14(6):1370. DOI: 10.3390/cancers14061370.
20. 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. (In eng). DOI: 10.3389/frai.2020.561802.
21. NHS England. AI Diagnostic Fund. (https://transform.england.nhs.uk/ai-lab/ai-lab-programmes/ai-in-imaging/ai-diagnostic-fund/).
22. Lawrence RD, E. Massou, E. Sherlaw-Johnson, C. Ramsay, A. Walton, H. O'Regan, T. Gleeson, F. Crellin, N. Herbert, K. Ng, P, L. Elphinstone, H. Mehta, R. Lloyd, J. Halliday, A. Morris, S. Fulop, N. Artificial Intelligence for diagnostics in radiology practice: A rapid systematic scoping review. In preparation.
23. Rapid Service Evaluation Team. Mixed-Method Evaluation of Implementing Artificial Intelligence (AI) in Chest Diagnostics for Lung Disease (Phase 1) - Summary of findings. ([to add]).
24. Ramsay AIG HK, Lawrence R, Sherlaw-Johnson C, Bagri S, Crellin N, Dodsworth E, Elphinstone H, Halliday A, Lloyd J, Massou E, Mehta R, Morris S, Ng PL, Walton H, Fulop NJ. Mixed-methods evaluation of implementing artificial intelligence in chest diagnostics for lung disease. Submitted (under review).
25. Best A, Greenhalgh T, Lewis S, Saul JE, Carroll S, Bitz J. Large-system transformation in health care: a realist review. The Milbank Quarterly 2012;90(3):421-456. DOI: 10.1111/j.1468-0009.2012.00670.x.
26. Turner S, Ramsay A, Perry C, et al. Lessons for major system change: centralization of stroke services in two metropolitan areas of England. Journal of health services research & policy 2016;21(3):156-165. DOI: 10.1177/1355819615626189.
27. Fulop NJ, Ramsay AI, Vindrola-Padros C, et al. Centralisation of specialist cancer surgery services in two areas of England: the RESPECT-21 mixed-methods evaluation. 2023.
28. Fulop NJ, Walton H, Crellin N, et al. A rapid mixed-methods evaluation of remote home monitoring models during the COVID-19 pandemic in England. Health and Social Care Delivery Research 2023;11(13):1-151. DOI: 10.3310/FVQW4410.
29. Walton H CC, Litchfield I, Sherlaw-Johnson C, Georghiou T, Massou E,tomini S, Sidhu M, Ellins J, Fulop NJ. . Exploring the use of the Major System Change framework to learn about rapid implementation of services during emergency contexts: A case study of COVID-19 remote home monitoring services. . Submitted (under review).
30. Walton H DM, Peter M, McInnes-Dean H, Mellis R, Allen S, Fulop NJ, Chitty LS, Hill M. Evaluating the implementation of rapid prenatal exome sequencing (pES) services in England. . Submitted (under review).
31. Faric N, Hinder S, Williams R, et al. Early Experiences of Integrating an Artificial Intelligence-Based Diagnostic Decision Support System into Radiology Settings: A Qualitative Study. Stud Health Technol Inform 2023;309:240-241. (In English). DOI: 10.3233/SHTI230787.
32. Rainey C, O'Regan T, Matthew J, et al. UK reporting radiographers' perceptions of AI in radiographic image interpretation - Current perspectives and future developments. Radiography (Lond) 2022;28(4):881-888. (In English). DOI: 10.1016/j.radi.2022.06.006.

33. Ng CT, Roslan SNA, Chng YH, et al. Singapore radiographers' perceptions and expectations of artificial intelligence - A qualitative study. J Med Imaging Radiat Sci 2022;53(4):554-563. (In English). DOI: 10.1016/j.jmir.2022.08.005.
34. Clarke V, Braun V. Using thematic analysis in counselling and psychotherapy research: A critical reflection. Couns Psychother Res 2018;18(2):107-110. (In English). DOI: 10.1002/capr.12165.
35. Neale J. Iterative categorization (IC): a systematic technique for analysing qualitative data. Addiction 2016;111(6):1096-1106. DOI: 10.1111/add.13314.
36. Vindrola-Padros C, Chisnall G, Cooper S, et al. Carrying Out Rapid Qualitative Research During a Pandemic: Emerging Lessons From COVID-19. Qual Health Res 2020;30(14):2192-2204. (In eng). DOI: 10.1177/1049732320951526.
37. The lung clinical expert group. The national optimal lung pathway. (http://content.smallerearthtech.co.uk/system/file_uploads/16086/original/National_Optimal_LUNG_Pathway_Aug_2017.pdf).

# 21 Appendices

Appendix 1. Study Gantt chart

Appendix 2. Compliance with RSET EDI Assessment tool

## Appendix 1. Project Gannt chart

| Phase 2 | Sep-24 | Oct-24 | Nov-24 | Dec-24 | Jan-25 | Feb-25 | Mar-25 | Apr-25 | May-25 | Jun-25 | Jul-25 | Aug-25 | Sep-25 | Oct-25 | Nov-25 | Dec-25 | Jan-26 | Feb-26 | Mar-26 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Develop protocol | | | | | | | | | | | | | | | | | | | |
| Key Stakeholder engagement | | | | | | | | | | | | | | | | | | | |
| Development of ethics materials and topic guides | | | | | | | | | | | | | | | | | | | |
| Peer, PPIE and funder review of protocol | | | | | | | | | | | | | | | | | | | |
| PPIE workshop- scoping group | | | | | | | | | | | | | | | | | | | |
| Identify and approach trusts | | | | | | | | | | | | | | | | | | | |
| Ethical approval | | | | | | | | | | | | | | | | | | | |
| Study sites approval | | | | | | | | | | | | | | | | | | | |
| Data collection | | | | | | | | | | | | | | | | | | | |
| Data analysis | | | | | | | | | | | | | | | | | | | |
| Formative feedback | | | | | | | | | | | | | | | | | | | |
| Write Up | | | | | | | | | | | | | | | | | | | |
| Submission of funders report | | | | | | | | | | | | | | | | | | | |
| Summative dissemination - publications etc. | | | | | | | | | | | | | | | | | | | |
| Advisory Panel Meetings | | | | | | | | | | | | | | | | | | | |

**Appendix 2. Equality, Diversity, and Inclusion (EDI) Assessment Tool**

| Stage of project (linked to flow chart) | Activity | Reviewed? | Notes on how this was considered within this project, and decisions made.<br>If activity not considered, please briefly add details on why this was not possible in this particular evaluation. |
|---|---|---|---|
| Building initial team | 1. Ensure evaluation teams include a diverse range of team members *[e.g. gender, age, ethnicity, seniority and other characteristics*] | YES | RSET team is diverse in terms of gender, age, ethnicity and seniority. This is reflected in this project team. |
| | 2. Ensure project steering groups include a diverse range of evidence users and healthcare professionals *[e.g. gender, age, ethnicity, seniority, role and other characteristics]* | YES | Good range of expertise/specialty, gender, ethnicity. |
| | 3. Ensure project PPIE panel includes a diverse range of patients/carers. *[e.g. gender, age, ethnicity, experience and other characteristics]* | YES | Range of gender, age, ethnicity among the public contributors. |
| Discovery and scoping | 4. Consult with **PPIE group** and **evidence users** (through scoping discussions) to understand EDI implications of both the intervention and our evaluation. | YES | Yes – this was a focus of discussions at<br>• initial meetings with PPIE members<br>• stakeholder workshops for Phase 1 and 2<br>• stakeholder engagement during scoping for Phase 1 and 2, including meetings and e-mail consultation<br>• public contributor on the advisory panel |
| | 5. During scoping conversations, the way in which PPIE members and evidence users are consulted should be adapted appropriately for each audience. *For example, it may be necessary to provide information in alternative formats other than standard text if people need or prefer that.* | YES | We asked PPIE workshops attendees about their preferences around sharing of information or opportunities to feed back. For Phase 2, these discussions also focused on approaches to recruiting patients and carers, and the structure and content of patient and carer interview topic guides. |
| | 6. Use EDI published frameworks (e.g. Health Inequalities Assessment Tool;[101] INCLUDE framework;[99] toolkit for increasing participation of Black, Asian and Minority Ethnic (BAME) groups in health and social care research).[102]<br><br>*These frameworks will help ensure that our projects are designed to be inclusive and address appropriate questions (e.g. considering underserved groups and wider protected characteristics, barriers to inclusion and steps to overcome barriers).* | YES | We used the Health Inequalities Assessment tool to ensure thorough consideration of EDI throughout all stages of the project. For example, this helped to identify potential EDI issues relevant to this topic including bias in the AI tool testing, trust and location related inequalities, and accuracy of AI for those with different characteristics or conditions. |

| | | | |
|---|---|---|---|
| Stakeholder engagement | 7. Discuss project with **project PPIE group and project advisory group** and ensure projects address EDI issues, including:<br>a. Whether and how different communities were involved in planning,<br>b. Whether and how research approaches accommodate and measure potential impact on EDI considerations,<br>c. Evaluating the intervention's impact on access, patient experience, engagement, and outcomes across different communities)<br>d. work with stakeholders to reflect on progress of the work and ensure our findings address implications for EDI. | PARTLY | We discussed project specific EDI issues with our PPIE members, Project Advisory Panel, stakeholder workshop, and RSET Stakeholder Advisory board. |
| Data collection – **focus** | 8. Develop research questions that address any issues of inequalities, inequities and disparities, as appropriate. | YES | Research questions consider implications of AI for EDI. |
| | 9. Identify how any relevant quantitative data reflects population diversity. | YES | Flowing from Phase 1, quantitative analyses will attempt to capture data in relation to population diversity where available. |
| Data collection – **site recruitment** | 10. Select study sites to represent a range of characteristics wherever possible (including geography, ethnicity, rurality, socioeconomic status). | YES | See sampling strategy, in terms of geographic location (including socio-deprivation, inequalities and clinical pathways). Final sample included majority of networks – good range of characteristics. |
| Data collection – **participant recruitment** | 11. Plan to recruit samples of patients, carers and staff that include a range of participants of different ages, gender, ethnicities, living circumstances, educational qualifications, work situations, and disability. | YES | See sampling: for our patient and carer interviews, we will seek to recruit patients across a range of characteristics , including health outcome following review of scan (and therefore care pathway), and factors relating to socio-demographic characteristics (e.g. gender, age, ethnicity, disability) |
| | 12. Where possible, compare our study sample characteristics to national or local populations accessing and delivering services (e.g. see[15]) | YES | Where available, we will be drawing on national datasets, e.g. HES and DIDS. |
| | 13. To support recruitment of a range of participants, consider the following strategies and other strategies as necessary (depending on appropriateness for each evaluation and conversations with stakeholders and PPIE panel):<br>a. Translating research materials into a range of languages or different formats where appropriate, e.g braille, or British sign language | Not applicable | We will be recruiting patients via hospital trusts/services, as we wish to interview people who have undergone diagnostic scans supported by AI. However, our sample has been designed to capture perspectives of less well served communities. |

| | | | |
|---|---|---|---|
| | b. Community outreach to recruit participants (e.g. through patient and staff organisations)<br>c. Offer different modes of data collection (e.g. in person, telephone or online for interviews/focus groups/observations and online or paper surveys),<br>d. Offer different options for participation (e.g. participant only, participant and carer, or carer only interviews)<br>e. Offer translation services to facilitate interviews.<br>f. Ensuring participants have reasonable access to participating in the study<br>*It may be helpful to look at the NIHR's definition of underserved communities when thinking about how best to recruit different groups* | | We will ask staff to recruit across a range of patient characteristics (see above).<br><br>We have capacity to translate our recruitment documentation in the event that potential interviewees require this, and we will explore the option of offering an interpreter service for interviews.<br><br>All recruitment documentation will be available in hard copies (if required), as well as electronically and compatible with screen reader software.<br><br>We will also offer potential interviewees the option of discussing the research with a member of the team in the event that they require further information. |
| Analysis | 14. Use frameworks to support equity-focused analysis where appropriate (e.g. EquIR).[59] | NO | We have not yet drawn on equity focused analysis frameworks, but is something that will be considered for upcoming analyses. |
| | 15. If available, analyse data to identify differences in service use and outcomes across different population groups | Not applicable | Where possible, we will do this in both our qualitative and quantitative workstreams. |
| | 16. Work with stakeholders **(project advisory group and PPIE)** to reflect on progress of the work and ensure our findings address implications for EDI | YES | EDI as a standing agenda item on weekly team meetings and advisory group meetings. |
| Dissemination | 17. Work with stakeholders **(project advisory group and PPIE)** to develop and agree a dissemination and mobilisation strategy that supports sharing findings with all relevant audiences (including diverse and underserved communities). | YES | Yes – have consulted dissemination strategy with wide range of stakeholders as part of peer review and meetings, including with third sector organisations and PPIE representatives. |
| | 18. Work closely with stakeholders **(PPIE panel, and project advisory group)** to share findings (e.g. as co-authors and co-presenters). | YES | We continue to work closely with stakeholders (PPIE members, project advisory group) to develop dissemination outputs. |
| | 19. If quantitative analyses of differences between population groups has not been possible, make recommendations about how to enable this for future evaluations. | YES | Again, we will explore this in our analyses (including our workshops with national and service-level stakeholders), and reflect on this in our resulting reports. |

*Note:* Throughout all our activities, we will be facilitated by guidance on effective EDI. [e.g. National Institute for Health and Care Research. Equality, Diversity and Inclusion Toolkit 2022. Retrieved 09/12/2022 from https://www.rdsresources.org.uk/edi-toolkit] / NIHR EDI strategy (2022-2027*) https://www.nihr.ac.uk/documents/equality-diversity-and-inclusion-strategy-2022-2027/31295*

Note: This tool was completed on 4th February 2025 during study protocol development phase. We will review the tool again during the data collection and analysis phase.