Mixed-method evaluation of implementing artificial intelligence in chest diagnostics for lung disease

Study protocol (v1.1, 12/06/2024)

Principal Investigator:

Dr Angus I G Ramsay (UCL)

Research team:

Chris Sherlaw-Johnson, Nuffield Trust

Emma Dodsworth, Nuffield Trust

Dr Nadia Crellin, Nuffield Trust

Dr Efthalia Massou, University of Cambridge

Dr Kevin Herbert, University of Cambridge

Prof Stephen Morris, University of Cambridge

Rachel Lawrence, UCL

Pei Li Ng, UCL

Holly Elphinstone, UCL

Dr Holly Walton, UCL

Raj Mehta, Public contributor

Emily Slade, Public contributor

Joanne Lloyd, Public contributor

Yasmin Rahman, Public contributor

Amanda Halliday, Public contributor

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.

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, and evidence suggests that AI can contribute to accurate diagnosis, reduce errors, and improve efficiency. 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.

In June 2023, NHS England announced the Artificial Intelligence Diagnostic Fund (AIDF), which is funding 11 networks of NHS Trusts across England to implement AI for chest diagnostics in 2024.

Aims and objectives

Our evaluation is the first phase of a planned two-phase evaluation. Our findings will both inform a Phase 2 evaluation and/or any future longer-term evaluations.

We will evaluate early deployment and implementation of AI for chest diagnostics as part of AIDF, to explore factors influencing implementation, and identify settings and data sources for a potential phase 2 evaluation and/or future longer-term evaluations.

Our research questions are:

1. How can we best collect patient and public perceptions of using AI diagnostic tools in clinical practice?
2. How can services best measure the impact of AI deployment on patients and the clinical pathway?
3. What are the key cost components of AI tools for chest diagnostics that are necessary for an economic evaluation of the AI diagnostic tools in clinical practice?
4. How are AI tools for chest diagnostics procured, deployed and implemented at network and trust levels?
5. What are stakeholder experiences (staff and AI suppliers) of the use of AI in chest diagnostics and associated care pathways?
6. Which factors influence implementation at network and trust levels? (including contextual factors and implications for EDI)

Methods

This will be a rapid, mixed-method evaluation of early deployment and implementation of AI for chest diagnostics, to be conducted over 10months to inform a potential phase 2 evaluation and/or future longer-term evaluations.

We will conduct a rapid scoping review followed by stakeholder consultation discussions (RQ1-3).

We will combine qualitative, quantitative, and health economic perspectives (RQ2-6). We will engage with network leadership of 11 networks, and we will conduct 3-4 in-depth case studies. We will use stakeholder interviews, non-participant observations

of oversight meetings, and analysis of relevant planning and progress documents, to: i) analyse deployment and implementation at network and trust levels including influential factors and stakeholder experiences and perspectives on early implementation (RQ4-6), ii) identify relevant outcomes, available data, and network/trust capability to collect and analyse these data, and provide advice on effective data use (RQ2) and iii) map chest diagnostic pathways, identify key costs and available data sources and tools, and explore whether it is possible to estimate costs related to different clinical pathways, including AI and non-AI pathways (RQ3).

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 four public members with an interest in chest diagnostics. All five attend team meetings and have supported the planning and writing of this protocol (e.g., commenting on drafts and contributing to planning discussions). They will be involved in writing any recruitment documents and research tools. 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 for members of the public with experiences of and interest in these services. Attendees supported our proposed approach and shaped the focus of our plans.

All PPIE involvement activities will be compensated in line with INVOLVE payment guidance.

The study will have an Advisory Group (see Section 11.3), which will include a range of stakeholders with relevant expertise, including patients, carers, and/or representatives of relevant charities.

Timelines for delivery

- *February 2024:* Protocol drafted, shared with peer reviewers and NIHR; developed research tools
- *February 2024:* Finalise protocol in light of peer review and NIHR feedback; obtain ethical and local permissions.
- *February-July 2024:* Scoping review and stakeholder consultation workshops (workstream 1)
- *March (after approvals)-August 2024:* data collection, rolling analysis, integration, and formative feedback (workstream 2-5)
- *July-November 2024:* complete project; share summative Phase 1 findings.

Anticipated impact and dissemination

We will share formative lessons on factors influencing implementation and potential ways to address challenges. We will also share summative lessons on how delivery, impact, and patient and public experiences of services might be monitored and evaluated.

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., the AIDF weekly network meeting and drop-in session), 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.

Through these, we anticipate addressing important gaps in the evidence base highlighted by the NICE evidence generation plan for AI in radiotherapy (published September 2023)[1] and influencing how implementation and impact of AI for chest diagnostics are monitored and evaluated at national, network, and Trust/service levels. We will also help shape the approaches taken in phase 2 and/or future evaluations, which will provide further important insights on progress and impact of AI for chest diagnostics.

PLAIN ENGLISH SUMMARY

Background

Chest x-rays and chest CT (computed tomography) scans help healthcare teams diagnose illnesses and make decisions about which treatment patients should get. Artificial intelligence (AI) describes computer systems that can be trained to recognise patterns in data and help solve problems. It is thought that AI might be able to help healthcare teams make faster and more accurate diagnoses, reduce pressure on healthcare teams, and reduce overall healthcare costs. However, we still need to know more about how AI can be used to look at chest x-rays or chest CT scans, including how it is put into action; the effect of AI on healthcare outcomes, resources, and costs; what healthcare teams think about using AI in their services; and what patients, carers, and the public think of AI for chest x-rays and chest CT scans.

In 2023, NHS England launched a pilot scheme to support introduction of AI for chest x-rays and chest CT scans at selected NHS hospitals across England, where AI is to support specialists in making treatment decisions. These services are expected to start using AI for chest x-rays and chest CT scans between June and October 2024.

What we aim to do

We aim to carry out a 10-month Phase 1 study of the early stages of putting AI for chest x-rays and chest CT scans into action, so we can:

1. Understand what helps with these processes or makes them more difficult, and
2. Find settings and sources of information for the next part of the study and further evaluations, which will be longer and more in-depth.

We will answer the following questions:

1. How can we best collect patient and public perceptions of using AI diagnostic tools in clinical practice?
2. How can services best measure the impact of AI deployment on patients and the clinical pathway?
3. How can we best measure the costs and resources involved in using AI tools for chest diagnostics?
4. How is AI being used to support analysis and reporting of chest x-rays and chest CT scans?
5. What do healthcare staff and AI suppliers think about these processes?
6. What helps and gets in the way of using AI to look at chest x-rays and chest CT scans? (including any impacts on equalities, diversity, and inclusion)?

How we will do this:

We will answer our questions in different ways.

We will look at previous research, to understand how best to study AI for chest x-rays and chest CT scans, in terms of impact on patients and care, costs, patient and carer experiences, and how to understand what the public think. We will also hold workshops to get views from members of the public.

We will speak with people involved in putting AI for chest x-rays and chest CT scans into action. We will observe meetings that discuss progress. We will look at documents covering the planning and progress of work at network and NHS Trust levels.

Patient and Public Involvement and Engagement

Patients and the public are central to this study. Our team includes the RSET patient & public involvement co-lead and four public members who are interested in AI for chest x-rays and chest CT scans. All five go to team meetings and have supported the planning and writing of this protocol (e.g., commenting on drafts and joining planning discussions). They will be involved in writing any information sheets and research tools. They will help us think through what our research means, and we will support them to be part of sharing our findings, (e.g., writing papers and presentations).

We discussed our plans at a workshop for members of the public with experiences of and interest in these services. Their feedback helped us to finalise our proposed approach and focus.

Our study will have an Advisory Group, which will include patients, carers, and/or the representatives of relevant charities.

We will pay people for their help and time in line with national guidance.

Sharing what we learn

We will share our findings as the study progresses. We will share our final lessons through academic journal articles, presentations to the national team, participating networks and services, and academic and professional conferences. We will produce accessible summaries of our findings, including slide sets, blogs, or animations.

Why is this research important?

AI for chest diagnostics may improve the speed, accuracy and efficiency of diagnosing serious chest conditions, so could result in important benefits for patients, the public, and healthcare services. It is also possible that AI implementation causes harm to some patients (potentially not yet identified) – remembering that if an algorithm makes a diagnostic error, it will repeat this error on every occasion. It is also important to check if the benefits of AI are increased or decreased when used in practice. It may also have an effect on training and skill mix, and this may have a delayed but important impact on care and costs. However, changing services is complex, and different approaches may work better (or worse) in different settings. Therefore, it is important to study how AI for chest diagnostics is put into action in a range of settings, to ensure that future plans to use AI for diagnostics consider relevant issues.

RESEARCH PROTOCOL

1. BACKGROUND AND RATIONALE

Need for AI within NHS context/policy

In England, the potential for digital technology to support the delivery of high quality and efficient care in the NHS has been highlighted in key healthcare policy documents.[2-4] Examples of digital technologies include tools to support patients with accessibility to care or healthcare services (e.g., healthcare apps, e-health consultation services, digital records), and tools to support staff (e.g., decision support tools and artificial intelligence).[2, 3] Additionally, current developments include using Artificial Intelligence (AI) to assist with automated tasks, diagnostic reporting, and image analysis (usually by being a second or concurrent reader (NHS refs 1-3) or depending on the pathway, using High Confidence Normals to report a normal or abnormal CT result).[1-3] In doing so, the integration of AI may help to reduce the workload burden on the NHS workforce to enable staff to have time for more complex and patient-facing tasks that AI could not perform.[2] This is particularly important in a UK health context, where the NHS faces challenges with workforce shortages and inconsistencies with demand vs resource[4]. Such inconsistences may be influenced by health inequalities, for example, geography and availability of local services, resulting in variations in access to care.[4] The implementation of AI also aligns with wider system goals to continually improve and advance the healthcare system.[2]

AI is described and understood as advanced technology that can perform complex tasks associated with human intelligence.[5-8] Machine learning is a subclass of AI[9] that is currently being explored for use within the NHS. Machine learning involves algorithms performing complex tasks by learning from patterns in the data.[3, 5, 9] This can be achieved by supervised learning, where the algorithm is specifically trained to interpret data, unsupervised learning where the algorithm interprets data without human input/training and reinforcement learning where the algorithm is able to self-learn and evolves as it interacts with the data.[3, 9]

Use of AI in radiology

Within the field of radiology, AI can be used to aid the early detection of cancer, for example by assisting with diagnostic reporting.[8, 10] When doing so, the purpose of an AI tool is to work alongside clinicians to detect suspected abnormalities and prioritise scans which may require a more urgent review.[1, 3] However, the NICE evidence generation plan for using AI to analyse chest X-Rays for suspected lung cancer (published September 2023) has highlighted a need for further evidence in several areas, including: time saving and resource use, adverse effects, performance in different patient groups, ease of use and perceived impacts.[1]

Effectiveness and cost-effectiveness

Research studies have investigated the use of AI to help assist with the detection of different types of cancer, such as breast,[11] head and neck[12] and lung.[5, 8] Focusing on lung cancer specifically in line with the scope of this project, research has studied clinical effectiveness and diagnostic accuracy. For example, a systematic review and meta-analysis evaluated the diagnostic value of AI for lung cancer and concluded that AI tools can accurately assist with detection, reduce and prevent errors, and improve overall efficiency.[13] Additional literature reports similar benefits to using AI for this purpose, particularly in relation to supporting clinical decision making and reducing

workforce burden (e.g. by assisting with repetitive tasks).[8, 9] Despite this, literature has reported limitations in the current evidence base, such as the data mainly being focused on selected cases/small samples and retrospective (8, 12, 5), with a lack of data demonstrating how it is used accurately as part of diagnostic pathways to impact patient care (5).

Implementation of, and experience with AI

Despite this, the integration of AI to support cancer diagnostics also raises concerns and challenges. Reviews have reported on issues regarding inconsistencies and variation in diagnostic accuracy,[14] bias in AI technologies[5, 6] and the need for an infrastructure and workforce capacity to facilitate implementation.[5, 7] When focusing on patient and public perceptions, issues can arise about the lack of personalisation, privacy and acceptability of, or confidence in the technology to provide accurate results.[10, 15-17] In line with policy aims and guidance on how to best use AI in diagnostic practice, evidence suggests that patients and the public often view AI as a tool to complement and assist clinical staff, but not replace them entirely.[10, 17-19]

Context for this study – Artificial Intelligence Diagnostic Fund

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.[20] The fund will focus on chest x-rays and chest CT scans to improve the diagnosis of lung cancer[20] and potentially help to address the current unmet need for faster chest x-ray reporting[20]. 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 .[20]

This study

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 these tools can and are being used in real-world settings.[1, 21] Further gaps in the evidence include research focused on the experiences and perceptions of implementing AI diagnostic tools (e.g., ease of use, confidence in and acceptability of AI), the cost implications in practice, and how AI may work for different groups.[1] For recommendations to be made regarding the implementation of AI diagnostic tools, these knowledge gaps need to be addressed.[1] Therefore, in this phase one study we will evaluate the early deployment of AI tools implemented as a result of the AIDF. Using a mixed-methods rapid approach, this study will evaluate the early deployment and implementation of AI tools for chest diagnostics and inform future evaluation of these tools both in the larger phase 2 study and other longer-term evaluations.

2. PROTOCOL DEVELOPMENT AND RESEARCH TEAM

2.1 Scoping phase

During our scoping phase and to develop this protocol, we reviewed relevant literature covering the general use of AI in healthcare and how it is implemented, perceptions on using AI in a health context (including barriers and facilitators), the use of AI for chest x-rays/CT and/or lung cancer specifically, inequalities relating to AI, and cost effectiveness. We engaged with relevant stakeholders to have ongoing discussions about the project, including academics (conducting work in the AI field), stakeholders involved in clinical practice (e.g., consultant radiologists and diagnostic radiographers), representatives from policy teams, occupational bodies (e.g., Royal

College of Radiologists, Society & College of Radiographers), representatives of patient groups, Voluntary Community or Social Enterprise (VCSE) organisations (e.g., the Patients Association), and stakeholders involved in regulation of AI within the NHS. In addition to this, we have been attending the weekly stakeholder network meetings, drop-in sessions and national evaluation subgroup meeting to engage with implementation updates and present/discuss our research plans to receive their feedback on the workstreams detailed in this protocol.

2.2 Research team

Researchers working in the NIHR RSET team (NJF, AR, CSJ, ED, NC, EM, KH, RL, PLN, SM, HE and HW) and Public Contributors (RM, ES, JL, AH and YR) will deliver the independent service evaluation. The team will work closely with national stakeholders and local teams (including implementation leads and clinicians); it will be overseen by an independent project advisory group including researchers, policy makers, the voluntary sector, and patient/carer representation.

3. AIMS AND OBJECTIVES

This project will evaluate early deployment of AIDF, to explore factors influencing implementation, and identify settings and data sources for Phase 2 and/or other further longer-term evaluation. Our objectives will be:

- To analyse local implementation and associated barriers and enablers at two levels: i) network level, and ii) trust level.
- Explore ways in which further longer-term evaluations might analyse implementation and measure the impact of AIDF at network and local levels

3.1 Research questions

Our research questions are:

1. How can we best collect patient and public perceptions of using AI diagnostic tools in clinical practice?
2. How can services best measure the impact of AI deployment on patients and the clinical pathway including implications for safety and health inequalities?
3. What are the key cost components of AI tools for chest diagnostics that are necessary for an economic evaluation of the tools?
4. How are AI tools for chest diagnostics procured, deployed and implemented at network and trust levels?
5. What are stakeholder experiences (staff and AI suppliers) of the use of AI in chest diagnostics and associated care pathways?
6. Which factors influence implementation at network and trust levels? (including contextual factors and implications for EDI)

For examples of sub-questions covered within each research question, see Appendix 1. See Figure 1 for a summary of the design.

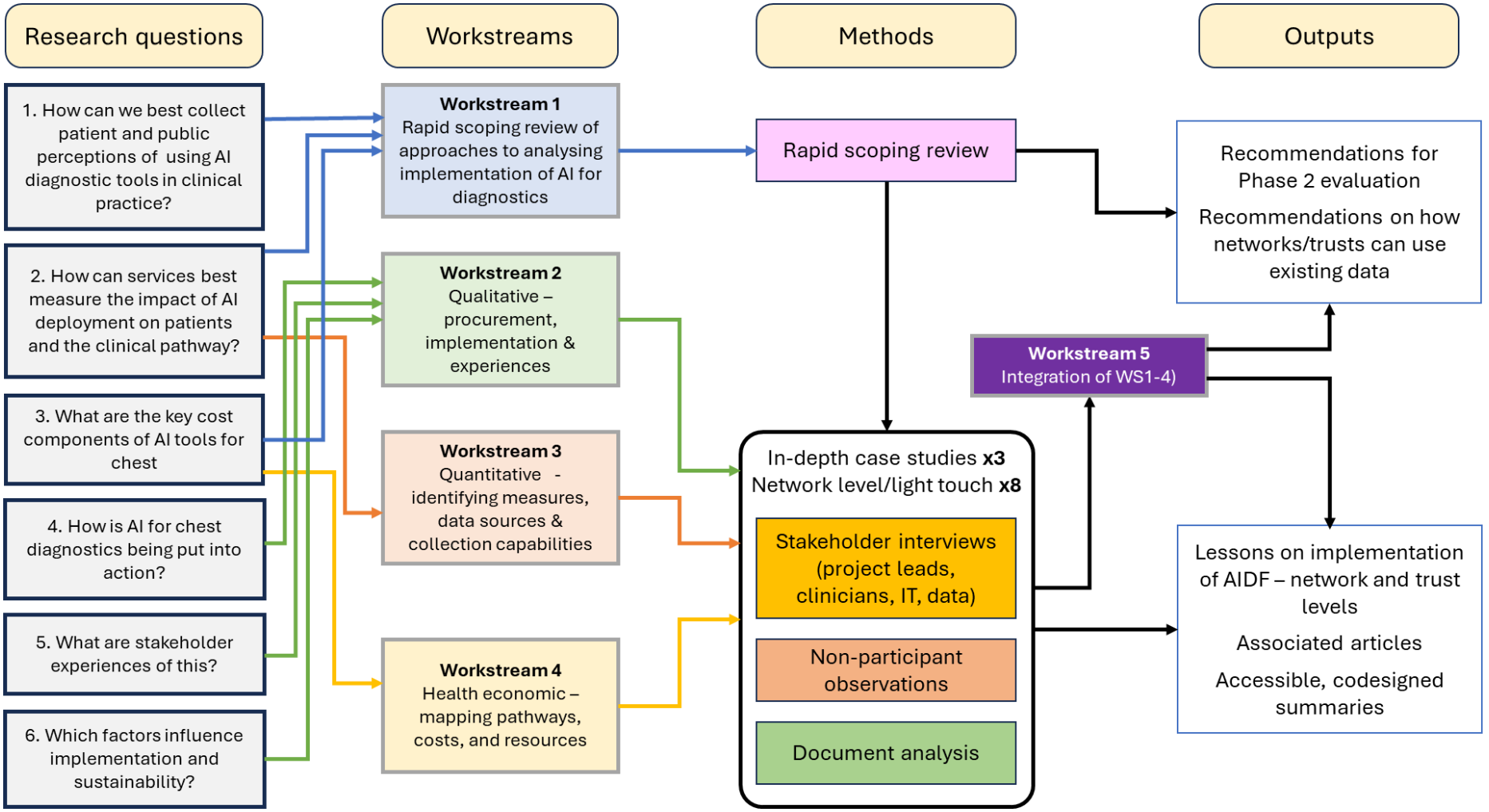


Figure 1. Overview of study design

4 STUDY DESIGN & METHODS OF DATA COLLECTION

4.1 Design

This is a proposed multi-site rapid study that will combine qualitative and quantitative perspectives to explore early deployment and implementation of Artificial Intelligence tools for chest diagnostics (chest x-ray and chest CT).

This evaluation will take place over 10 months (February to November 2024) and will inform the development of the Phase 2 evaluation and/or further longer-term evaluations.

4.2 Theoretical framework

The study will be guided by the Non-adoption, Abandonment, and challenges to the Scale-up, Spread and Sustainability (NASSS) of technologies for health and social care framework.[22] This framework describes barriers/ facilitators that may influence the implementation of technological innovations and considerations for evaluation[22] and is therefore a suitable guide for the present study. The NASSS framework will be used to assist with the development of interview topic guides and the interpretation of study findings.

4.3 Methods

This rapid evaluation has five workstreams (see Figure 1).

Workstream 1. Rapid scoping review with stakeholder consultation discussions (RQ1-3)

This workstream aims to conduct a review of the implementation and evaluation of AI tools for diagnostics. The review will explore: 1) How AI tools for diagnostics are perceived and experienced by patients, carers, and members of the public and how this evidence has been collected and analysed, 2) What quantitative evidence on outcomes exists and using what measures and 3) What economic evidence exists and what are the cost components that have been used.

This workstream will be led by Rachel Lawrence with contributions from other team members, including representatives from each workstream with qualitative, quantitative, and economic expertise.

Methods:

To meet this aim and inform methods of data collection for Phase 2 and/or future evaluations, we will conduct a rapid scoping review of the literature followed by stakeholder workshops to explore the relevance of review findings.

To ensure this method was appropriate, we have conducted a very rapid initial scope of the literature using PubMed, which included search terms related to AI (e.g., artificial intelligence, machine learning, deep learning), diagnostics (e.g., diagnosis, diagnose), experience and/or perceptions (e.g., view, experience, perspective, thoughts), quantitative outcome measures (e.g., benefits, accuracy, sensitivity), cost (e.g., cost-effectiveness, cost-benefit, health economics) and clinical and/or practice implementation (e.g., implement, deploy, adopt[33]). In line with how the review will be conducted, search terms were provided by representatives from each workstream. Results indicate relevant and sufficient literature which may be eligible, for example,

at least 10 studies have focused on patient perceptions of using AI for diagnostics, in addition to more than 12 studies that focused on the economic evaluation of AI implementation for diagnostic purposes). With many (e.g., 6 patient/public/carer studies, and 6 economic evaluations) studies focusing on AI for diagnosis in radiology, there is scope for us to focus the review further to include the topic of radiology relevant for AIDF.

With AI for diagnostics being a rapidly evolving field, this scoping review can add value to the existing literature by providing evidence on the implementation and evaluation of AI for diagnostics. Furthermore, the review has a broad focus (inclusive of grey literature and/or policy documents) with the aim of bringing together evidence related to methods of data collection for patient/carer/public perspectives and experiences, quantitative and cost-related outcome measures. The findings of this scoping review will be used to 1) inform components of future evaluation in relation to how data is collected for a Phase 2 evaluation in the context of AIDF; 2) feed WS2 of the current evaluation in terms of identifying recruitment and data collection approaches, additional questions, and frameworks to support analysis and interpretation; feed WS3 and WS4 in terms of relevant metrics and necessary evidence about outcomes, cost components, and cost model parameters.

The stages of the review are presented below.

Stage 1: Rapid scoping review

We will follow a rapid review method proposed by Tricco et al.[23] This method is systematic, but with adaptations to reduce the time required to carry out the review. We will use the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement.

The rapid scoping review of the literature will:

1. Identify the methods and approaches used to best support collection and analysis of data on patient, carer and public perceptions and/or experiences of AI as implemented for diagnostics.
2. Identify the type of questions that should be included when asking patients, carers and the public about AI as implemented for diagnostics.
3. Explore what is known about public, patient, and carer perceptions and experiences of AI when implemented for diagnostics for radiology
4. Identify outcome measures that have been used to evaluate the impact of AI for diagnostic imaging (on e.g., care delivery, and EDI), the quantitative data that have been analysed and commentary around their use.
5. Identify the key cost and resource components of procuring, deploying, and implementing AI diagnostic tools.

Eligibility criteria

Our review will focus on the implementation and evaluation of AI for diagnostics, including the methods and approaches used to collect data on patients, carers, and public perceptions and experiences of AI for diagnostics and studies carrying out quantitative assessments of the impact of these services on outcomes and costs.

Inclusion criteria

To be included in the scoping review, studies will need to (i) focus on the use and implementation of AI for diagnostics purposes. (ii) report on qualitative evidence which explores patient, carer and patient perceptions and/or experiences of AI as implemented for diagnostics, or provide a quantitative assessment of these services or cost components; (iii) be written in English and (iv) have been published in the last 5 years (since 2019, due to the rapid advancements of technology within healthcare during this period, and the publication of policy documents referring to AI in healthcare[2]). Peer reviewed published literature, grey literature and policy documents will be included. Depending on numbers of studies we find reporting quantitative findings we may restrict to studies within actual healthcare pathways whether as a clinical trial or in routine use.

Exclusion criteria

Studies: (i) focusing on the use of AI within healthcare but not specific to diagnostics in radiology, (ii) not written in English, and (iii) published prior to 2019 will be excluded.

Search strategy

A phased approach[23] will be used, which means that terms will be added based on keywords used within relevant published reviews on this topic. We will include search terms specific to AI (e.g., artificial intelligence, AI, deep learning, machine learning), diagnostic purposes (e.g., diagnostics/ diagnosis – if focusing on radiology additional search terms may include radiology/medical imaging/clinical imaging/CT/MRI) and patient, carer and public experiences or outcomes (e.g., view/ experience/perspective/thoughts, benefits/accuracy/sensitivity, cost-effectiveness /cost-benefits/health economics). To identify relevant papers, we will search the following databases: PubMed, Medline, Cinahl, PsycInfo and web of science. Where appropriate, we will hand search relevant journals. We will conduct a google search (including relevant policy websites) to obtain grey literature and policy documents.

One researcher will conduct the database search, input the records, and remove any duplicates.

Study selection

Studies will be screened based on the inclusion and exclusion criteria. The following phases will be used to screen studies identified during the search: (i) title, (ii) abstract and/or executive summary and (iii) full text. One researcher will screen the articles with additional researchers checking a percentage of articles in abstract and full text stages, in line with recommendations for conducting rapid reviews.[23]

Data extraction

Data extraction will be carried out using a form developed in Excel. This will relate to study characteristics and methods (e.g., title, date of publication, setting, aim, design, population and analysis) and findings from the reviews (e.g., patient, carer and public perceptions/experiences; quantitative metrics, data, outcomes, and challenges with data; and cost/resource components. This will be developed following initial screening

of articles and piloted by two researchers using a random sample of articles.[23] Any disagreements will be discussed until a consensus is reached.

Data synthesis

Data will be exported and we will descriptively synthesise study characteristics, including participant demographics and study setting/context. In line with the exploratory research question aiming to inform Phase 2, the findings will be synthesised using narrative synthesis.[24] Findings will be grouped into (i) the methods and/or approaches used for data collection and analysis, (ii) the types of questions asked and (iii) the key findings. These will be reported across all workstreams of qualitative, quantitative and cost.

Stage 2: Stakeholder workshops

The inclusion of workshops is consistent with recommended scoping review methods.[25]

The purpose of the stakeholder consultations with members of the public and AI users will be to explore their perspectives on findings from the rapid scoping review (e.g. the best methods/approaches to use, questions that should be asked and what outcome measures may be important for future studies).

Sampling and recruitment

Up to two online workshops will be conducted. For the first workshop, we will aim to recruit members of the public (aged 18+). For the second workshop, we will aim to recruit individuals (e.g. healthcare professionals, such as radiographer, radiologist, ED/hospital/primary care staff with access to PACS in their role) using AI for chest diagnostics. No data will be collected from patients during this Phase 1 review, but findings will inform how this may be achieved in Phase 2 and/or future evaluations.

Eligibility criteria: To be able to take part in the stakeholder consultation workshops, participants will need to be:

- Aged 18+
- Have the capacity to provide their informed consent
- Able to participate in an online group discussion
- A member of the public (for workshop 1) or a user of AI in practice for chest diagnostics (for workshop 2)

Participants will not be eligible to take part if they are under the age of 18 or are unable to participate/consent to being involved in an online group discussion.

Recruitment: Participants will be recruited mainly through social media, local charities or relevant organisations, and direct invitation. If interested, participants will contact the researcher and then be sent an information sheet and consent form. They will have at least 48 hours to review this and if still interested will be asked to provide informed consent (written consent – scanned forms or typewritten/electronic signature, or audio-recorded verbal consent). Following this, a date and time will be arranged for the workshop and the information will be sent to participants via email, including the link to join the online call.

Data collection and analysis

In advance of the workshops, participants will be sent a summary of preliminary scoping review findings. Prior to the start of discussions, participants will be provided with contextual information about the research project, aims and methods. Focus group consultations will follow a semi-structured topic guide, including introductory discussions and exploring their views on each part of the review findings. Vignettes describing case specific examples may be used to facilitate discussions. The workshops will be audio recorded (subject to consent being given), 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. Findings will be analysed using thematic analysis and an inductive approach to coding, drawing on the overarching theoretical framework guiding this protocol where relevant. The findings from the workshops will be integrated with review findings.

Workstream 2. Qualitative study of implementation and experiences (RQ4-6)

Aims

This workstream aims to evaluate and explore the early deployment and implementation of AI tools for chest diagnostics (chest x-Ray and chest CT), including i) reasons for AI deployment, ii) how AI tools have been procured and implemented at network and trust levels (including factors which influence early deployment) and iii) stakeholder perceptions and experiences of early deployment.

Note: The methods described in this workstream will also be used by Workstreams 3 and 4, addressing the overarching aim of making recommendations that can inform future implementation and evaluation.

This workstream will be led by Angus Ramsay with contributions from other team members.

Sampling

To evaluate early deployment and implementation as well as stakeholder experiences on both a wider network and trust level, we will sample networks at 2 levels (see Figure 2 for a summary):

- (i) Network only: all networks participating with AIDF for chest diagnostics (or as many as possible). This will involve:
    - Interviews at each network with one or two network representatives
    - Interview with the AI supplier
- (ii) In-depth case studies in three network areas, obtaining perspectives from network-level and two trusts within these areas. These will involve:
    - Interviews with members of staff per NHS Trust (including radiographer, radiologist, IT/data lead, Information governance leads/representatives)
    - Observations of relevant meetings
    - Analysis of relevant documentation

Eligibility criteria: To be eligible to take part in the interviews (workstreams 2-4), participants will need to meet one of the descriptions below:

- *Network leads:* Working as network lead in one of the 11 networks implementing the AI tools as part of AIDF.
- *AI suppliers:* representative of a supplier chosen to provide the AI tools for networks as part of AIDF.
- *Network team members:* Working within one of the selected network teams and involved in assisting with the implementation the AI tools as part of AIDF.
- *Trust level staff:* Working within one of the trusts selected who are involved with implementing the AI tools as part of AIDF. These roles may include radiologist, radiographer, Picture Archiving and Communication System (PACS) staff, and information governance leads/representatives.

All participants (across all levels of recruitment) will be over the age of 18 and have capacity to provide informed consent to participate in an online interview.

Participants will not be eligible to take part if they do not meet our eligibility criteria specified above, are under 18 or unable to participate in an interview/do not consent to take part.

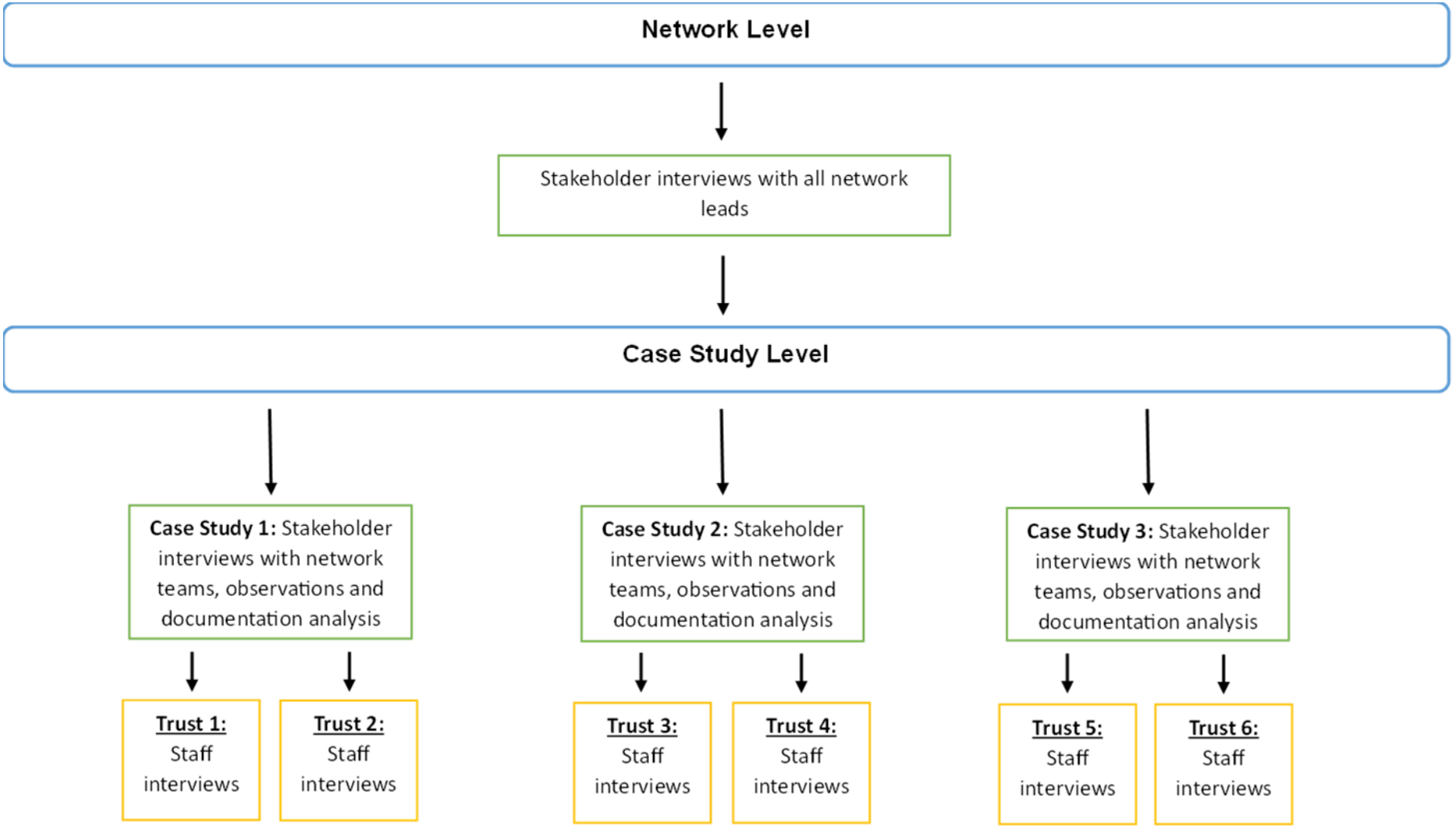


Figure 2. Summary of network level and case-study level data collection approaches

Methods –used across both levels

At network level and case-study level, we will use the following data collection methods: semi-structured interviews with stakeholders involved in procurement, deployment and implementation of AI diagnostic tools, documentary analysis of key documents and observations of relevant meetings. These will be described in further detail in the following sections.

Methods: Network level

We will conduct one semi-structured interview with a network lead/network leads for each network (*N* = up to 11) (see Table 1). The interviews will focus on obtaining a network perspective on procurement, development and how the AI tools are being implemented across and within different trusts. Network leads will be recruited by sharing study information in meetings and via email with all networks and with consent, receiving email contacts for potential participants.

At a network level, we will also interview the supplier providing the AI tools within the networks. Of the suppliers being used by networks (*N* = up to 3), we will aim to explore their perspectives on the process of procurement, early deployment and implementation, including how this may differ across the different networks involved. AI suppliers will be identified and recruited via email by the researcher.

Potential interviewees will be sent an information sheet and consent form. They will be given at least 48 hours to review and ask questions (guided by the individual's preferences). If the participant/s then agree to take part, they will be asked to provide informed consent (written – scanned forms/typewritten/electronic signature, or audio-recorded verbal consent). The researcher will then arrange a date and time for the interview to be carried out over the telephone or using an online platform such as Zoom or Microsoft Teams.

Table 1. Purposive sampling framework for interviews

Level	**Organisation**	**Interviews per organisation[4]**	**Number of interviews**
Network Teams[1]	11	1[3]	11
AI suppliers	3	1	3
Case study Trusts[2]	6	4[5]	24
Total:			**38**

Note: Numbers are maximums, not targets.
[1.] We aim to recruit all 11 networks.
[2.] We aim to recruit up to six NHS trusts, i.e., two in each of our three case studies.
[3.] We will aim to interview one person for each network, though some people may wish to take part in a joint interview with another colleague.
[4.] These interviews include topics from workstreams 2-4. Where appropriate additional ad hoc or follow up interviews may be undertaken for workstreams 3-4.
[5] Relevant roles include: radiologists, radiographers, IT staff, data staff

Methods: Case study sites (including trust level)

Three networks will be selected for in-depth case study work. We will aim to sample a range of networks and trusts which will be selected based on characteristics such as: Diagnostic type (CT or chest x-ray), the purpose/function of the AI (as specified in bids), geographical location (taking into consideration socio-deprivation, inequalities and clinical pathways), the size of the network/trust, the stage of deployment/ implementation and maturity of local data collection processes (see Table 2).

Interviews

The case study semi-structured interviews carried out with staff members will follow the sampling framework outlined in Table 1. In addition to the interviews with a network lead and AI supplier, we will also conduct additional interviews in two trusts per case-study network. We will aim to interview staff members (*N* = 3-4 per trust) involved in directly implementing the AI tools in clinical practice or those involved in supporting the implementation (e.g., this may include radiologists, radiographers, IT and Picture Archiving and Communication System (PACS) staff – these may differ between trusts) and information governance representatives. Potential participants (as identified by leads and trust teams) will be contacted by the researcher via email and sent an information sheet and consent form and given at least 48 hours to review and ask questions. If the participant/s then agree to take part, they will be asked to provide informed consent (written consent – scanned form/typewritten/electronic signature or audio-recorded verbal consent). The researcher will then arrange a date and time for the interview to be carried out via telephone or using an online platform such as Zoom or Microsoft Teams.

Table 2. Sampling characteristics

Variables	Description
Diagnostic type	Chest CT or chest x-ray
AI supplier	Range of AI suppliers
Geographical location	Where network and/or trusts are located across the country, ensuring spread across NHS in England
Stage of implementation	Which stage they are at in relation to implementation (e.g., how long they have been implementing the tools since funding and procurement)
Maturity of local data collection processes	Differences in trusts' abilities to collect appropriate data and to link patient records between datasets – indicator how Trusts are approaching collection of local benefits data

Interview topic guides will be developed for network leads, clinical staff, and AI suppliers. These will broadly cover the following topics: their role in AI, their understanding of AI, the process of implementation (including transparency), reasons for deployment, expectations of AI, barriers and facilitators to implementation,

approaches to monitoring uptake, outcomes, and safety, and reflections/learning on early and future implementation. The interviews will last between 30 and 45 minutes. Interviews may be shorter or longer than this depending on how much the interviewee would like to say in response to the questions. Prompts may be used to obtain further detail when needed. The interviews will be semi-structured, audio recorded (subject to consent being given), transcribed verbatim by a professional transcription service, anonymised, and stored in compliance with the General Data Protection Regulation (GDPR) 2018 and Data Protection Act 2018. Interview data will be collected alongside meeting notes and the documents provided by networks and trusts.

Documentary analysis

We will request key documents from stakeholders at a network and trust level (e.g., service specifications, bid documentation, and equality and health assessments – if available) and analyse these to provide insights into procurement, deployment, and ongoing implementation, to show how the use of AI tools is progressing in practice.

Observations

We will identify relevant planning and oversight meetings at network and trust level (e.g. service evaluation sub-group meetings, weekly AIDF network meetings, and the weekly network drop-in meetings, local planning and implementation meetings) and request permission to observe these. Before the meeting takes place, those attending will be made aware of our team's presence and information sheets about the study and purpose of the observations will be shared in advance. The chair of each meeting will be asked to provide consent for the formal observations to take place. We will ensure that all attendees know that personal information will not be reported in any outputs. If at any stage personal information is discussed in the meeting, we can temporarily leave and dial back in as guided by the lead chairing the meeting.

Data analysis

Data collection and analysis will be carried out in parallel. Emerging data (including interviews, documentary analysis and observations) will be collated using Rapid Assessment Procedures (RAP) sheets.[26] RAP sheets will be developed per network, case study site and trust, to facilitate cross-case comparisons and per population (to make comparisons between sub-groups). The categories used in the RAP sheets will be based on the questions included in the interview topic guide, maintaining flexibility to add categories which may also be informed by the observations and documentation analysis as the study progresses. Findings will be grouped into themes and sub-themes, supplemented with illustrative quotes.

Workstream 3. Quantitative (RQ2)

Aims

In this workstream we aim to investigate the quantitative data-related issues around enabling effective evaluation of AI tools for chest x-rays and chest CT scan, including:

I. Identifying relevant outcomes for a Phase 2 evaluation or other future evaluations of the technology and for ongoing monitoring of safety and effectiveness, including for an assessment of potential inequalities.
II. Identifying any data that can be used to assess these outcomes, understand data quality and identify any gaps where data doesn't currently exist.
III. Evaluate the capabilities of sites and networks regarding:
    a. Data capture
    b. Linkage to other data systems
    c. Quantitative assessment of potential inequalities
IV. Undertake analyses of data already available throughout the project, as appropriate.
V. Advise on how sites/networks can use their data effectively, such as facilitating evaluation, working within existing constraints and monitoring of effectiveness and safety.

This workstream will be led by Chris Sherlaw-Johnson with contributions from other team members.

Methods

The activities undertaken within this workstream will be exploratory and based on interviews, review of literature, documents and data supplied by networks for NHSE, and investigation of existing data sources.

Scoping review *(workstream 1)*

As outlined in *workstream 1*, we will conduct a scoping review to identify relevant outcome measures and any specific issues related to data collection for evaluating AI deployments.

Documentary analysis

We will also analyse documents (including those obtained within *workstream 2* as well as reports from national bodies such as NICE and the MHRA). We will also review the periodic returns from Networks populating the NHSE Benefits Register.

Observations of key meetings

Alongside colleagues in *Workstream 2* we will attend and observe key meetings with stakeholders.

Interviews at network and NHS Trust levels

Some information will be gathered from interviews with trusts and networks conducted for *workstream 2* to obtain experiences of data collection as part of the AI deployment. This will include how they monitor safety and effectiveness, and the barriers to measuring outcomes required by NHSE for their assessment of benefits. Relevant questions will be added to the topic guides developed within W*orkstream 2*. Particular topics covered include:

- How the AI is being used by trusts/networks.
- What data collected by trusts and its quality.
- Trust's processes for data collection.
- What is required for trusts/networks to populate measures in the benefits register.
- Their ability to collect data for the outcome measures we identify.

As described in *workstream 2*, Table 1, additional ad hoc interviews may also take place with individuals at trust level who have particular knowledge about the software and local data collection processes. These interviews may be audio-recorded (with consent) if necessary (see *workstream 2*, 'methods' section for further details).

We also propose to obtain information from suppliers on AI training data and their adaptability to intervention sites. Specifically, asking them for their assessment of any similarities between the product training data and the population the product is implemented on and any concerns if this data differs. We will also ascertain whether they update the product to improve performance at sites. However, we recognise that this may not be possible in every case due to intellectual property issues, for example, and have identified this as a risk in Table 3, Section 12.3.

Sample

The site and interview sample will be drawn from that selected for *workstream 2* (see Table 2). For workstream 3 purposes, one of the interviewees per site will cover data aspects (see Table 1). These individuals will be identified through initial discussions with the trust or network. We also recognise that these may lead to further ad hoc interviews with more individuals in order to follow-up specific details.

Data analysis

Due to the timescale of the project and the timeline of AI deployment across sites, we do not anticipate we would be undertaking any major analysis of data to assess impact. If appropriate and the data is available in time, we would analyse data from the Benefits Register as collected by trusts. Specifically, we will assess data completeness and variation of outcomes between trusts. This will help us to identify challenges trusts are facing in data collection and early indication of whether benefits are being realised. Since benefits metrics are collated by NHSE, we will liaise with NHSE to ensure work is not being duplicated. Benefits metrics data is aggregated at trust level and does not include patient data.

Investigation of existing data sources

We will identify relevant data sources, for example, the national lung cancer audit and cancer registries, to explore what is currently being collected and assess how these might assist with evaluation of the AI tools. Where appropriate, we would advise on sources that trusts and networks could use for baseline data in further evaluation.

Outputs

The findings from our investigations will feed into a list of recommendations that will help trusts and networks with data capture and ensure that their services are evaluable across all relevant domains. Building on and complementing existing frameworks, we will also advise on a dataset for longer-term evaluation, collection processes, and analyses, where relevant.

As mentioned above, if we analyse aggregated data provided for the NHSE Benefits Register we will provide early assessments of the impact of AI deployment on care processes.

Workstream 4. Cost study (RQ3)

Aims

Given the recognised evidence gaps in the real-world costs associated with AI use,[1] in the Phase 1 evaluation we will explore and define economic aspects relating to both current practice and the use of AI tools in chest diagnostic pathways (chest x-ray and chest CT) for lung cancer diagnosis, including:

i. Identification of key cost components related to the deployment of AI tools
ii. Rapid review of currently available data sources to inform (i)
iii. The adaptation of previously published, or the development of *de novo* cost-related data collection tools

The identification of all relevant cost components will provide a framework for data collection and cost estimation of AI tool deployment forming a robust basis for future full economic evaluations.

This workstream will be led by Kevin Herbert with contributions from other team members.

Methods

Sample

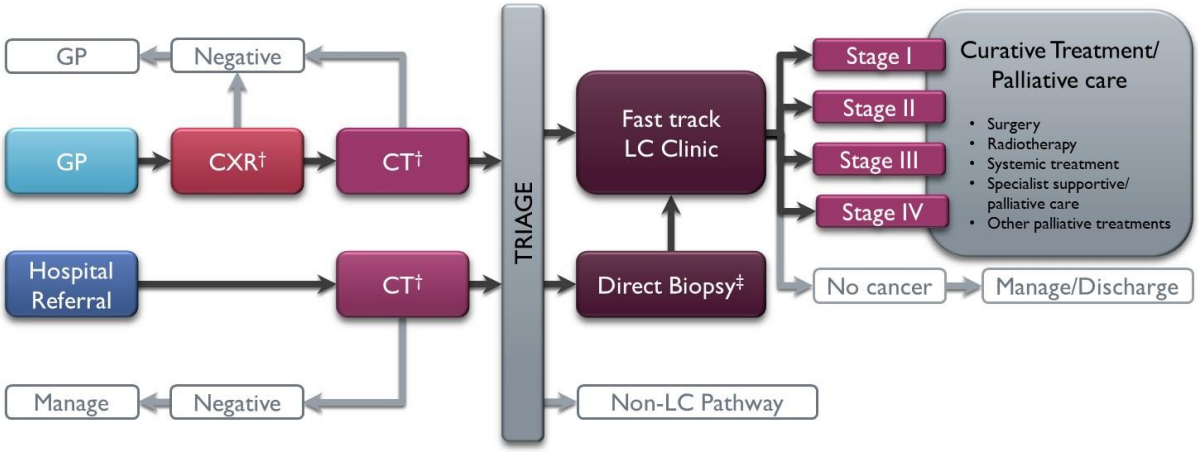
The cost study will focus on lung cancer diagnostic services (AI and non-AI) across the three NHS Networks and NHS Trusts sampled under *workstream 2* (see Table 1 and Table 2).

Measures

Resource use data not routinely captured by the trust will be sought using a collection tool (either adapted from published literature, or specifically developed for this study) to be completed by individual sites and from answers to scoping and follow-up questions, in collaboration with service delivery and clinical experts (e.g., manual results validation, conflict resolution process, establishment of in-house auditing/safety monitoring processes, post-diagnosis patient treatment pathways). If required, data from previously published peer-reviewed studies may also be used. The data collection tool will include a short consent process at the beginning.

Data collection

Participants will be asked to complete the consent section at the start of the data collection tool. Identification of data to be collected and of their respective sources, will cover those relevant to AI tool deployment and service delivery, and will be mapped with respect to the National Optimal Lung Cancer Pathway (Figure 3). Data to inform the cost model which is not available in the those routinely collected by the trust (e.g., details of diagnostic and patient treatment pathways, estimates of relevant input parameters etc.), will be sought via questionnaire, in consultation with clinical experts, or from previously published studies (where required). Where in consultation with clinical experts, these data will be sought through follow-up questions in *ad hoc* interviews.



† 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).
Derived from National Optimal Lung Cancer Pathway, NHS England, 2020. https://www.cancerresearchuk.org/sites/default/files/national_optimal_lung_pathway_aug_2017.pdf
Abbreviations: CT, ; CXR, chest x-ray; GP, general practice; LC, lung cancer.

Figure 3. Overview of lung cancer diagnostic pathway.

Data analysis

To inform future economic modelling studies, the cost study report will describe the pre-, during- and post-chest imaging pathways, outlining the key data to be collected, relevant collection points and sources for resource use valuation (where required).

Workstream 5. Integration of data collection and findings

This workstream will be led by Angus Ramsay with contributions from other team members.

As described above, the evaluation workstreams draw heavily on the same methods, i.e., interviews, observations, documentary analysis, and evidence review. Therefore, the qualitative, quantitative, and health economic teams, and our PPIE members have collaborated closely in developing the study design for each workstream.

Research tools (interview topic guides, observation frameworks, evidence extraction tools) will be codesigned by the different workstream teams to address the priorities of all workstreams: this will ensure efficient data collection, minimising burden on participating organisations and individuals.

We will organise our analysis and findings from all workstreams around our research questions, whilst also drawing on the NASSS framework[22] which addresses factors such as implementation approaches and resources involved in monitoring impact of change.

The team will meet on a weekly basis to discuss progress and/or cross-cutting learning, as appropriate over the course of the study. Additionally, regular workstream-specific meetings will take place, involving input qualitative, quantitative, and health economic teams.

5 RECRUITMENT AND CONSENT FOR PARTICIPANTS IN EMPIRICAL RESEARCH

Workstream 1: To recruit members of the public and/or chest diagnostic AI users to stakeholder consultation workshops, we will share study information on social media, via third sector organisations or relevant organisations, and send direct invitations to participate. After receiving expressions of interest, we will send information sheets (which will explain the aims of the discussion events, length of the events and details on data collection and data storage, and who to contact should any questions or problems occur and details about participant withdrawal) and consent forms. We will give participants at least 48 hours to review these documents. Participants will be informed that taking part is voluntary, findings will be fully anonymised and kept confidential, and that they are free to withdraw at any point. Participants will also have the opportunity to ask questions at any time. If participants want to take part after reading the information sheet and asking any questions, 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 to ensure informed and voluntary participation.

The stakeholder consultation workshops will be carried out via an online platform (Zoom or Microsoft Teams). As the discussions will take place remotely, we will also check with the participants if they are still happy to consent to take part at the start of the discussion.

Workstreams 2-4: We will engage with teams prior to the data collection phase to increase awareness of the work and explore best ways to inform potential participants about the evaluation. When recruiting staff members for interview, we will email eligible team members (at network and trust level) to ask if anyone would be interested in taking part. For those who express their interest, we will send information sheets (which will explain the purpose of the interview, length of interviews and details on data collection and storage). We will give participants at least 48 hours for these documents to be reviewed. Information sheets will also include details about who to contact should any questions or problems occur and details about participant withdrawal. Participants will be informed that taking part is voluntary, findings will be fully anonymised and kept confidential, and that they are free to withdraw at any point. Participants will also have the opportunity to ask questions at any time.

If participants want to take part after reading the information sheet and asking any questions, 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 to ensure informed and voluntary participation.

Interviews will be carried out via telephone or an online platform (Zoom or Microsoft Teams). As interviews will take place remotely, we will also check with the participants if they are still happy to take part at the start of the interview.

For meeting observations, we will liaise with the chair of the meeting regarding whether observation will be possible and appropriate. We will send the meeting chair the information sheet (which will explain the purpose of the observation, details on data collection and storage, who to contact with questions or any problems) and consent form. If the chair is happy in principle for the meeting to be observed, they will be asked to provide written consent. At the start of each meeting, we will also gain verbal

consent from meeting attendees for the study team to observe and take anonymised notes.

For *workstream 4,* we will ask participants to complete a consent section at the start of the cost data collection tool.

6 PATIENT AND PUBLIC INVOLVEMENT AND ENGAGEMENT

Patients and the public have shaped the study from its inception. The involvement of diverse perspectives is not only integral to ethical research but also ensures that the study remains relevant and impactful.

Prior to commencing the study, RSET PPIE Co-Leads (PLN & RM) advertised the RSET PPIE Panel role on the NIHR People in Research platform. Subsequently, a thorough application shortlisting process was undertaken, inclusive interviews were conducted with PPIE representatives, and collaborative decision-making was employed. Information about the study was then shared via email with the RSET PPIE Panel, inviting expressions of interest in joining the study team. The selection process prioritised diversity, resulting in the inclusion of four public contributors as full members of the study team: Joanne Lloyd, Emily Slade, Amanda Halliday and Yasmin Rahman. Together with PPIE Co-Lead Raj Mehta, they have actively engaged in the weekly project team meetings since the project's inception.

Feedback and suggestions from the public contributors have been diligently recorded on meeting agendas. They have suggested valuable insights, such as proposing scoping questions for stakeholders (e.g., is AI looking at the right things, or does it just look like it is looking at the right things?) and commenting on study planning documents (e.g., including an explanation about data storage and GDPR/Data Protection). These public contributors will continue to contribute to all aspects of the study, including design, recruitment documents, topic guide for interviews, interpretation of findings, co-authoring articles and summaries, and other dissemination activities.

In addition, we held a PPIE Scoping workshop virtually on 17th January 2024 to discuss the study scope, research questions and designs with service users. Attendees, recruited from a poll of over 30 service users, particularly those with lung conditions or experience with chest x-ray/CT scan and/or their family members. The attendees (N=6) were highly supportive of the proposed work, noting the importance of issues around communication, transparency, and trust. They also suggested potentially valuable approaches to collecting data in Phase 2. A similar PPI workshop may be arranged in the future during the interpretation of findings and the design of dissemination outputs, ensuring ongoing collaboration with the wider public.

The Evaluation Advisory Group (Section 11.3) is expected to include patients, carers or representatives from the voluntary sector. The study has allocated a budget for patient contributions in line with good practices identified by NIHR INVOLVE[27], ensuring fair compensation and recognising the value of PPIE.

In summary, the research team works in a highly inclusive and collaborative manner, valuing diverse expertise and perspectives, including those of the public, patients, and carers. This builds and sustains mutually respectful and beneficial relationships.

7 EQUALITY, DIVERSITY AND INCLUSION (EDI)

To ensure that our project thoroughly and comprehensively considers issues of equality, diversity and inclusion, we will apply our NIHR Rapid Service Evaluation Team (RSET) EDI project-specific checklist (see Appendix 3) at two stages during this project: (i) during development of the project evaluation and (ii) following data collection and analysis. The checklist 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).

8 STUDY SCHEDULE

The proposed timeline for the study is presented below. See Appendix 2 for a detailed Gantt chart.

- *AI tool for chest diagnostics implemented in practice: rolling implementation* - February to October 2024
- *Study design & develop protocol*: January 2024
- *Development of ethics materials and topic guides*: January 2024
- *Peer review of protocol*: February 2024 (returned 9th Feb 2024)
- *Protocol review by NIHR*: February 2024 (returned 9th Feb 2024)
- *PPIE review of protocol:* February 2024 (returned 9th Feb 2024)
- *Set up study advisory group*: January 2024
- *Workstream 1 Conduct systematic review*: February-July2024
- *Ethics approval*: February 2024
- *Local approvals:* MarchJune 2024
- *Workstream 1 - Stakeholder consultation workshops:* July 2024
- *Workstream 2-4 Data collection: Ma*rch 2024 (as soon as local approval obtained) - August 2024
- *Workstream 2-4 Data analysis:* May 2024-October 2024
- *Workstream 5 Integration of findings:* May 2024-October 2024
- *Write up:* July 2024-November 2024
- *Submission of final report:* November 2024
- *Summative dissemination:* August 2024 onward

9 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).

10 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, RL, HW, NJF.

10.1 Data management

Data will be managed in line with legal and regulatory requirements, including the General Data Protection Regulation (GDPR) and the Data Protection Act (2018), and necessary research approvals. Dr Angus Ramsay will act as the data controller for this study. He will process, store, and dispose of all data in accordance with all applicable legal and regulatory requirements, including the General Data Protection Regulation (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.

A Data Sharing and Processing Agreement will be in place between the research team and NHS England for the purposes and duration of this evaluation.

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 (University College London and Nuffield Trust) and/or quantitative researchers (from Nuffield Trust) and health economists (from University of Cambridge) where appropriate topics regarding data and costs are discussed.

Participant consent forms and audio-recordings of interviews and workshops 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 stored securely in locked filing cabinets within the UCL office. Any participant consent forms sent via email will be uploaded directly to the UCL DSH.

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 have access. Only the research team will have access to participants' personal data (i.e., name and contact details). A password protected spreadsheet of interviewees and their contact details will also be held on the UCL DSH. Participant identifier codes will be stored in the UCL DSH and kept separate from study data.

Workstream 2: Trust and/or Network-level benefits registers will be accessed from a protected area of the FutureNHS website 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.

Workstream 3: Lung cancer diagnostics and outcome data (aggregated across the deployment 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**.**

11 PEER AND REGULATORY REVIEW

11.1 Ethics

Based on the Health Research Authority (HRA) decision tool and consultation with the UCL/UCLH Joint Research Office, this study is classified as a service evaluation. We will submit a low-risk research application to the UCL Research Ethics Committee.

Although this is a low-risk evaluation, we are aware of the sensitive nature of this work for organisations and individuals. 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.

11.2 Management

This project is led by Ramsay 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.

11.3 Governance

Independent oversight and advice will be provided in following ways. First, 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 on a regular basis.

12.4 Peer review and quality assurance

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

As noted above, the project will be discussed in monthly RSET Executive Management meetings, a key focus of which is quality assurance around research design and analysis.

12 ASSESSMENT AND MANAGEMENT OF RISKS

12.1 Ethical considerations

During interviews and workshops, we will be asking staff, AI suppliers and members of the public 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.

12.2 Management issues

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.

12.3 Risks and mitigation

Potential risks are highlighted in Table 3.

Table 3. Potential risks and mitigation strategies

Workstream	Risk	Impact	Likelihood	Mitigation
All	Limited engagement from networks	Limited range of experiences captured	Medium	Active engagement at national and network levels
All	Limited engagement from Trusts/ interviewees	Limited range of experiences captured	Medium	Active engagement at national and network levels
WS2-4	Limited engagement from AI suppliers	Limited range of experiences captured	Medium	Team to engage suppliers now that they are in the public domain and provide assurances about anonymisation and confidentiality
WS2-4	Delays in obtaining Data Sharing and Processing agreement	Unable to access quant data and supporting documents	Medium	Commenced process in December, currently sitting with UCL contracts
WS2-4	Delays in obtaining local research governance permissions	Delay to collecting interview, observation, and documentary data	Medium	Engaging with local teams to begin process as early as possible. Working with national leadership to explore potential to emphasise need to prioritise rapid processing of requests. We will also have a letter of support from NHS England that may support prioritisation of approval requests.

13 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].

14 MONITORING AND AUDITING

The project lead (AR) 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 11.3).

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).

15 TRAINING

The project lead (AR) together with the chief investigator of NIHR RSET (NJF) will review and provide assurances of the training and experience of all staff working on this study.

16 INSURANCE

University College London 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. University College London 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.

17 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 (AR) 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.

18 PUBLICATION AND DISSEMINATION

This study's rapid methods will enable identification of formative and summative lessons on the procurement, deployment, and implementation of AI for chest diagnostics (and potentially diagnostics in other settings), factors influencing implementation, and potential ways to address challenges. We should therefore help support ongoing implementation of AI for chest diagnosis. This study will provide learning on potential data sources for analyses of how AI affects delivery, outcomes, and cost-effectiveness of these services, and how patient, carer, and public perspectives might be analysed. The study therefore has potential to influence current and future implementation of AI for chest diagnostics (and potentially other diagnostics), address key gaps in the evidence base, and shape future approaches to monitoring and evaluation of AI in chest diagnostics. A key output of this evaluation will be an evaluation framework to guide subsequent phase 2 and/or future longer-term evaluations.

Methods of dissemination will be discussed and agreed with stakeholders (e.g., our PPIE panel, project advisory group, NHS England, and other key stakeholders) at an early stage. Examples of dissemination outputs could include: (i) the sharing of lessons via presentations at national and network level meetings (e.g., the AIDF weekly network meeting and/or drop-in session), other meetings with staff from trusts involved in the AIDF, and via NHS Futures platform, (ii) the production of evaluation summary slide set or report and (iii) accessible summaries such as blogs or animations.

To maximize the impact of this learning, we will also produce academic journal articles and present findings in a range of academic and professional-focused conferences.

19 REFERENCES

1. 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. Retrived 30/01/24 from: Overview | Artificial intelligence-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals: early value assessment | Guidance | NICE.
2. NHS England. The Long-Term Plan 2019. Retrieved 30.01.2024 from https://www.longtermplan.nhs.uk/publication/nhs-long-term-plan/
3. NHS England. Artificial Intelligence and machine learning 2023. Retrieved 30/01/2024 from https://www.england.nhs.uk/long-read/artificial-intelligence-ai-and-machine-learning/#background
4. NHS England. The Long-Term Workforce Plan. 2023. Retrieved 30/01/2024 from https://www.england.nhs.uk/long-read/nhs-long-term-workforce-plan-2/
5. Gleeson F, Revel M-P, Biederer J, Larici AR, Martini K, Frauenfelder T, et al. Implementation of artificial intelligence in thoracic imaging—a what, how, and why guide from the European Society of Thoracic Imaging (ESTI). European Radiology. 2023;33(7):1-10. doi: 10.1007/s00330-023-09409-2
6. Harwich E, Laycock K. Thinking on its own: AI in the NHS. Reform Research Trust. 2018.
7. Istasy P, Lee WS, Iansavichene A, Upshur R, Gyawali B, Burkell J, et al. The impact of artificial intelligence on health equity in oncology: Scoping review. Journal of Medical Internet Research. 2022;24(11):e39748. doi: 10.2196/39748
8. Mathew CJ, David AM, Mathew CMJ. Artificial intelligence and its future potential in lung cancer screening. EXCLI journal. 2020;19:1552. doi: 10.17179/excli2020-3095
9. Chiu H-Y, Chao H-S, Chen Y-M. Application of artificial intelligence in lung cancer. Cancers. 2022;14(6):1370. doi: 10.3390/cancers14061370
10. 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 Journal. 2021;27(2):14604582211011215. doi: 10.1177/14604582211011215
11. Pacilè S, Lopez J, Chone P, Bertinotti T, Grouin JM, Fillard P. Improving breast cancer detection accuracy of mammography with the concurrent use of an artificial intelligence tool. Radiology: Artificial Intelligence. 2020;2(6):e190208. doi: 10.1148/ryai.2020190208
12. Mahmood H, Shaban M, Rajpoot N, Khurram SA. Artificial Intelligence-based methods in head and neck cancer diagnosis: an overview. British journal of cancer. 2021;124(12):1934-40. doi: 10.1038/s41416-021-01386-x
13. Liu M, Wu J, Wang N, Zhang X, Bai Y, Guo J, et al. The value of artificial intelligence in the diagnosis of lung cancer: A systematic review and meta-analysis. Plos one. 2023;18(3):e0273445. doi: 10.1371/journal.pone.0273445
14. Fletcher RR, Nakeshimana A, Olubeko O. Addressing fairness, bias, and appropriate use of artificial intelligence and machine learning in global health. Front Artif Intell. 2021;3:561802. doi: 10.3389/frai.2020.561802
15. Adams SJ, Tang R, Babyn P. Patient perspectives and priorities regarding artificial intelligence in radiology: opportunities for patient-centered radiology. Journal of the American College of Radiology. 2020;17(8):1034-6. doi: 10.1016/j.jacr.2020.01.007

16. Fritsch SJ, Blankenheim A, Wahl A, Hetfeld P, Maassen O, Deffge S, et al. Attitudes and perception of artificial intelligence in healthcare: A cross-sectional survey among patients. Digital health. 2022;8:20552076221116772. doi: Artn 20552076221116772

10.1177/20552076221116772
17. Ibba S, Tancredi C, Fantesini A, Cellina M, Presta R, Montanari R, et al. How do patients perceive the AI-radiologists interaction? Results of a survey on 2,119 responders. European Journal of Radiology. 2023:110917.
18. Hemphill S, Jackson K, Bradley S, Bhartia B. The implementation of artificial intelligence in radiology: a narrative review of patient perspectives. Future Healthc J. 2023;10(1):63-8. doi: 10.7861/fhj.2022-0097
19. Musbahi O, Syed L, Le Feuvre P, Cobb J, Jones G. Public patient views of artificial intelligence in healthcare: A nominal group technique study. Digital Health. 2021;7:20552076211063682. doi: 10.1177/20552076211063682
20. NHS England. AI Diagnostic Fund 2023. Retrieved 30/01/2024 from https://transform.england.nhs.uk/ai-lab/ai-lab-programmes/ai-in-imaging/ai-diagnostic-fund/
21. Strohm L, Hehakaya C, Ranschaert ER, Boon WP, Moors EH. Implementation of artificial intelligence (AI) applications in radiology: hindering and facilitating factors. European radiology. 2020;30(10):5525-32. doi: 10.1007/s00330-020-06946-y
22. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, Hinder S, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. Journal of medical Internet research. 2017;19(11):e8775. doi: 10.2196/jmir.8775
23. Tricco AC, Langlois E, Straus SE, Organization WH. Rapid reviews to strengthen health policy and systems: a practical guide: World Health Organization; 2017.
24. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct of narrative synthesis in systematic reviews. A product from the ESRC methods programme Version. 2006;1(1):b92.
25. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International journal of social research methodology. 2005;8(1):19-32.
26. Vindrola-Padros C, Chisnall G, Cooper S, Dowrick A, Djellouli N, Symmons SM, et al. Carrying out rapid qualitative research during a pandemic: emerging lessons from COVID-19. Qualitative health research. 2020;30(14):2192-204. doi: 10.1177/1049732320951526
27. National Institute for Health and Care Research. Payment guidance for researchers and professionals 2023. Retrieved 30/01/2024 from https://www.nihr.ac.uk/documents/payment-guidance-for-researchers-and-professionals/27392
28. For equity. Health inequalities assessment toolkit 2022. Retrieved 22/12/2022 from https://forequity.uk/hiat/
29. National Institute for Health and Care Research. Improving inclusion of under-served groups in clinical research: Guidance from INCLUDE project 2022. Retrieved 23/12/2022 from https://www.nihr.ac.uk/documents/improving-inclusion-of-under-served-groups-in-clinical-research-guidance-from-include-project/25435

30. NIHR Applied Research Collaboration East Midlands. Toolkit for: Increasing participation of Black Asian and Minority Ethnic (BAME) groups in health and social care research 2018. Retrieved 23/12/2022 from https://arc-em.nihr.ac.uk/clahrcs-store/increasing-participation-black-asian-and-minority-ethnic-bame-groups-health-and-social
31. Fulop N, Walton H, Crellin N, Georghiou T, Herlitz L, Litchfield I, et al. A rapid mixed methods evaluation of remote home monitoring models during the COVID-19 pandemic in England Health Soc Care Deliv Res. 2022. doi: 10.3310/hsdr-tr-132703
32. Eslava-Schmalbach J, Garzón-Orjuela N, Elias V, Reveiz L, Tran N, Langlois E. Conceptual framework of equity-focused implementation research for health programs (EquIR). Int J Equity Health. 2019;18(1):1-11. doi: 10.1186/s12939-019-0984-4
33. Yin, J., Ngiam, K. Y., & Teo, H. H. (2021). Role of artificial intelligence applications in real-life clinical practice: systematic review. *Journal of medical Internet research*, *23*(4), e25759. DOI: 10.2196/25759.

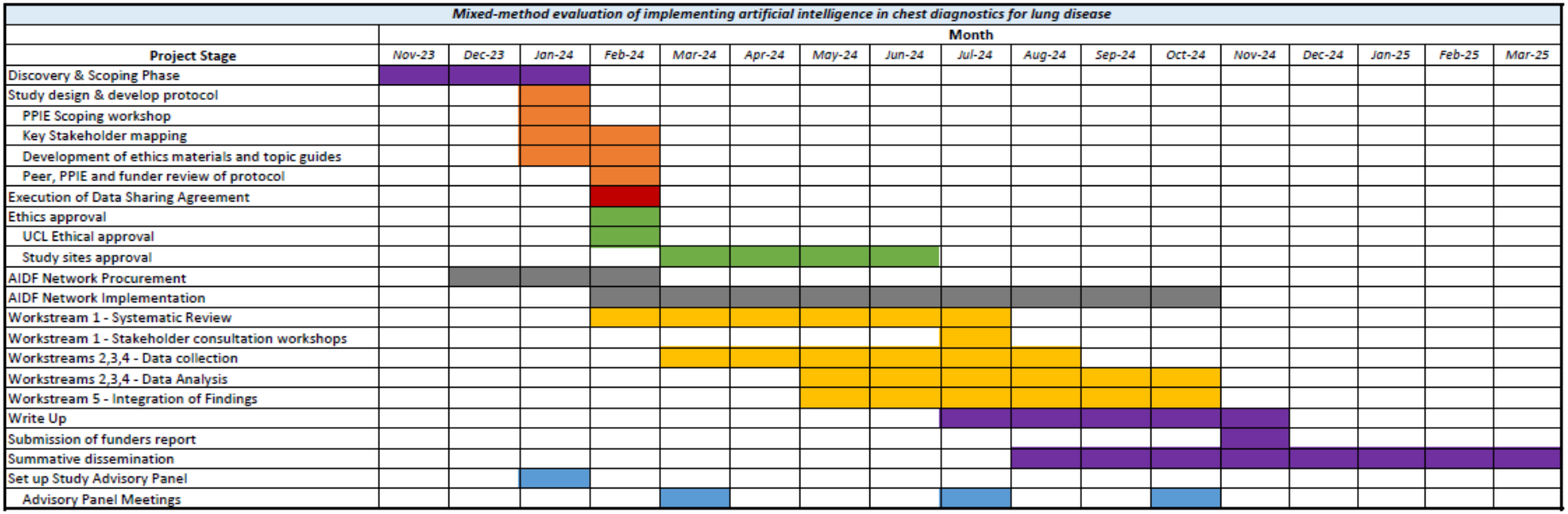
20 APPENDIX

Appendix 1. Example sub-questions to be answered within each research question

Research question	Example sub questions	Evidence gap reported in NICE[1] evidence generation plan that this workstream relates to
1. How can we best collect patient and public perceptions of using AI diagnostic tools in clinical practice?	- What methods and approaches have been used in published literature to collect and analyse data on patient, carer, and public perceptions/experiences of AI? - What type of questions have been asked in previous studies? - What is known about perceptions and experiences for diagnostics for radiology in published literature? - What do members of the public and AI users think of these findings and the methods that might be appropriate in future research? - Implications for future evaluations?	Informing future evaluations of evidence gaps relating to: - patient experience
2. How can services best measure the impact of AI deployment on patients and the clinical pathway?	- What outcome measures are needed for a full evaluation of AI and for ongoing monitoring of safety, effectiveness and potential inequalities? - How are services monitoring safety? - What quantitative data exists that can be used to assess these outcomes? What is the data quality? - Where are there gaps in current available data? - What are the capabilities of sites and networks regarding data collection and extraction? - Implications for future evaluations?	Informing future evaluations of evidence gaps including: - referrals to CT scan, - time to chest X-ray review, - CT referral and diagnosis, - diagnostic accuracy and technical failure rates, - evidence in populations with underlying conditions that could yield images difficult to interpret
3. What are the key cost components of AI tools for chest diagnostics that are necessary for an economic evaluation of the tools?	- What cost and resource components have been identified for deploying and implementing AI diagnostic tools in published literature? - What are the key cost components relating to deployment of AI tools? - What diagnostic and treatment costs are required for AI and non AI pathways? - Differences between costs of AI and non-AI based pathways? - Implications for future evaluations?	Informing future evaluations of evidence gaps including: - Software impact on healthcare costs and resource use -
4. How are AI tools for chest diagnostics procured, deployed and implemented at network and trust levels?	- What are the reasons for AI deployment? - How have AI tools been procured and implemented at network and trust levels? Any contrats with more conventional procurement approaches?	Informing evaluation of evidence gaps, including: - Clinician experience of using AI-derived software

	- Composition and approach of implementation team? - How did teams draw on evidence behind AI tech to inform procurement, pathway design, and monitoring plans? - Perceived incentives for implementation - Influence on care delivery – e.g. do clinicians still do the full review despite AI (i.e. increasing effort) - Intended vs actual use? - Similarities and differences between networks and within networks? - Unintended consequences? - Recommendations to improve implementation? - Implications for future evaluations?	- Evidence in populations with underlying conditions that could yield images that are challenging to interpret - Implementation considerations
5. What are stakeholder experiences (staff, and AI suppliers) of the use of AI in chest diagnostics and associated care pathways?	- What are staff members views of using AI? - What are staff members understanding of the use of AI? - What are staff members expectations of AI? - What are the barriers and facilitators to engagement with AI? - What are the barriers and facilitators to delivery? - Similarities and differences between networks and within networks? - What are staff's recommendations? - Implications for future evaluations?	
6. Which factors influence implementation at network and trust level? (including contextual factors and implications for EDI)	- Barriers and facilitators to procurement? - Barriers and facilitators to deployment? - Barriers and facilitators to implementation? (e.g. capacity, IT systems, IG processes) - Similarities and differences between networks and within networks? - Implications for EDI? - Implications for future evaluations?	

Appendix 2. Study Gantt chart – updated June 2024



Appendix 3. RSET project specific Equality, Diversity and Inclusion checklist

Stage of project (linked to flow chart)	**Activity**	**Done?**	**Notes on how this was considered within this project, and decisions made.** **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 AI 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 • Need to fill gap around voluntary sector rep
	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 on panel
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 • initial meetings with PPIE • stakeholder workshop • stakeholder meetings during scoping phase
	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 attendees for any preferences around sharing of information or opportunities to feed back.
	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] *These frameworks will help ensure that our projects are designed to be inclusive and address appropriate questions (e.g. considering underserved groups and wider*	YES	Using the ***Health Inequalities Assessment Tool*** to guide how we consider EDI during all phases of the study. The tool has five sections to be considered throughout study design. The toolkit provides a description about the purpose of each of the five sections, a reflection question and involve question, to help researchers think about EDI and how it can be considered when conducting the study. **Section 1:** Mapping health inequalities relevant to your research • **Purpose:** All research has dimensions of inequalities – need to map those most relevant to your project. Existing

	protected characteristics, barriers to inclusion and steps to overcome barriers).		literature will help but also engaging with people who have lived experience and/or people working in the field. • **Reflect:** Which dimensions of social and health inequalities are relevant to your research? *Examples:* Computational experts not having input from clinicians, who the AI tool has been tested on (often different to populations subsequently used on – usually tested on white males), trust and/or location related inequalities (e.g., different budgets, different training, socioeconomic differences, rural vs urban populations, differing attendance to cancer screening etc), we don't know enough about accuracy and how well AI tools work for different groups of people, physical characteristics which can make the scans harder to read (e.g., obesity) and how smoking can be more common in lower income groups (to be considered when looking at location of AIDF implementation). • **Involve:** How can people with relevant lived experience and/or policy and practice expertise help you to identify dimensions of inequalities relevant to your study? *Examples:* PPIE involvement throughout where EDI has been discussed (one PPIE workshop completed – more to follow), stakeholder engagement during scoping phase of the study which is ongoing – in all meetings we have discussed potential EDI considerations (advisory group has also been developed). We will also be conducting stakeholder consultations to discuss rapid review findings – specifically in relation to relevance and application. Stakeholder engagement has and will continue to involve a variety of different perspectives, from academic experts to clinicians and professional bodies, charities and third sector organisations. • *Anything else we could be doing?* Discuss inequalities with sites during recruitment phase? Reflect on EDI throughout recruitment and data collection/analysis? Consider providing EDI specific guidance for Phase 2?

			Section 2: Integrating an intersectional equity lens into research questions • **Purpose:** Encourages you (as researchers) to consider how you integrate dimensions of inequalities in your research questions. • **Reflect:** Will your research enable you to identify potential inequalities in social and health outcomes and explore structural causes of these? We have proposed to explore different ways to collect data for patients and members of the public through reviewing the available literature – which will then inform how this could be achieved in Phase 2, where further EDI considerations could be explored. Findings will be discussed through stakeholder consultations (as described above). Across workstreams, we aim to collect data from all network leads in the different locations implementing the AI tools, focusing on four case studies in more depth. We will aim for these case studies to be diverse in relation to sampling (e.g., considering their location, CT, or X-Ray etc). This, alongside meeting observations and documentation analysis, will help enable us to explore structural and contextual components of implementing the AI tools as part of AIDF (e.g., can think about how policy and processes may differ across sites). • **Involve:** How can people with lived experience and those with policy and practice expertise help you to embed relevant dimensions of inequalities into your research questions? PPIE involvement and stakeholder engagement has influenced the research questions and design. PPIE members are part of the research team and have provided feedback regularly. We have also presented proposed research questions to stakeholders during meetings to receive their feedback and in a PPIE advisory group, with inequalities being a discussion point of these meetings. • *Anything else we could be doing?* Ask staff about inequalities in the interviews? Have a discussion in the stakeholder consultations about EDI considerations when exploring patient and public experiences? Ask staff about

			training and resource in the interviews? Collect contextual information about sites and their populations if feasible? **Section 3:** Designing and conducting research sensitive to inequalities • **Purpose:** Design is often determined by the problem and/or questions which need to be addressed. Researchers need an understanding of what is known about inequalities in research area to help inform the design. Scoping work has been used throughout (meeting notes) to inform research design. PPIE team members have been involved in the proposed design of the study and provided feedback on a regular basis. • **Reflect:** Will your study design, including analytical techniques, enable you to explore differential impacts/experiences? We will use a mixed-methods approach which will explore a variety of different impacts/experiences. For example, as part of the qualitative workstream we are aiming to interview network leads, radiologists, radiographers, IT staff and AI suppliers across case study sites. This will ensure we can include staff in different roles and different levels of implementation (e.g., AI suppliers, network leads and then those using the AI tools in practice). In the review conducted, we will explore how to collect data from patients and members of the public to inform how this may be done in Phase 2. With this being a scoping review, we will aim to report on published literature and grey literature to review a wide range of evidence. Relevance of findings will also be discussed in stakeholder consultations – which we will aim to include a variety of different participants (e.g., members of the public, experts in the relevant field). • **Involve:** How can people with relevant lived experience and those with policy and practice expertise help you design and conduct research? Same comments above in relation to stakeholder engagement and PPIE team members (including the advisory group).

			• *Anything else we could be doing?* Links to above when considering EDI in the topic guides (including workshop discussions) and contextual information collected about case study sites – also EDI considerations in the analysis (something to think about moving forward)? **Information on Section 4 and 5 (relevant for later stages in the project):** **Section 4:** Prioritising findings relevant to inequalities in reporting and dissemination • Consider how to best report findings which relate to inequalities • Think about how to best disseminate – clear and accessible messages, directed at relevant audiences. Try to maximise impact. PPIE involvement important to ensure suitable. **Section 5:** Principles for research sensitive to intersectional inequalities • Researchers can work to remove biases (e.g., through reflection) • Research can be conducted in ways that challenge inequities – main focus of the tool: which is why it is important to reflect on throughout. • Research practices can positively promote greater equity – ensure inclusive approaches to the research
Protocol drafting and stakeholder engagement	7. Discuss project with **project PPIE group and project advisory group** and ensure projects address EDI issues, including: a. Whether and how different communities were involved in planning, b. Whether and how research approaches accommodate and measure potential impact on EDI considerations,	PARTLY	In process • EDI issues discussed with PPIE and at stakeholder workshop • Project advisory group yet to meet (though have explored some of these issues in stakeholder consultation discussions)

	c. Evaluating the intervention's impact on access, patient experience, engagement, and outcomes across different communities) d. work with stakeholders to reflect on progress of the work and ensure our findings address implications for EDI.		
Protocol drafting and data collection – **focus**	8. Develop research questions that address any issues of inequalities, inequities and disparities, as appropriate.	YES	RQs consider implications of AI for EDI
	9. Identify how any relevant quantitative data reflects population diversity.	YES	A focus of WS 1, 3, and 4 is how quant data might reflect population diversity
Protocol drafting & data collection – **site recruitment**	10. Select study sites to represent a range of characteristics wherever possible (including geography, ethnicity, rurality, socioeconomic status).	YES	Yes – see sampling strategy, in terms of geographic location (including socio-deprivation, inequalities and clinical pathways)
Protocol drafting & 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	Staff-focused data collection – can monitor diversity of participants
	12. Where possible, compare our study sample characteristics to national or local populations accessing and delivering services (e.g. see[15])	YES	Staff-focused data collection – can monitor diversity of participants
	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): a. Translating research materials into a range of languages or different formats where appropriate, e.g braille, or British sign language b. Community outreach to recruit participants (e.g. through patient and staff organisations) c. Offer different modes of data collection (e.g. in person, telephone or online for interviews/focus groups/observations and online or paper surveys),	Not applicable	In the main, this will be staff-focused data collection However, we will consider this for public-facing workshops

	d. Offer different options for participation (e.g. participant only, participant and carer, or carer only interviews) e. Offer translation services to facilitate interviews. f. Ensuring participants have reasonable access to participating in the study *It may be helpful to look at the NIHR's definition of underserved communities when thinking about how best to recruit different groups*		
Protocol drafting & analysis - analysis	14. Use frameworks to support equity-focused analysis where appropriate (e.g. EquIR).[59]	YES	TO DO: will be useful in analysing implications for EDI.
	15. If available, analyse data to identify differences in service use and outcomes across different population groups	Not applicable	Highly important for Phase 2 – Phase 1 will be making recommendations on this
	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 agenda item on weekly team meetings and PAG meetings
Protocol drafting & 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 shared dissemination strategy with wide range of stakeholders as part of peer review, and discussed options at e.g. Network meetings.
	18. Work closely with stakeholders **(PPIE panel, and project advisory group)** to share findings (e.g. as co-authors and co-presenters).	YES	TO DO - but clearly stated in protocol
	19. If quantitative analyses of differences between population groups has not been possible, make recommendations about how to enable this for future evaluations.	YES	Part of WS3&4 priorities in protocol
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*			