

Project INSIGHT: Improving diagnosis and treatment for lung cancer in London

Improving outcomes for Londoners, reducing inequalities and using NHS resources more effectively.

London Driver Project 2026-27

Project Leads

Richard Lee (RMNHSFT), Derralynn Hughes (NCLCA), Alex Blakoe (NCLCA), James Teo (KCL), Richard Dobson (KCL), Brendan Delaney (ICL), Céire Costelloe (ICL), Sunnia Gupta (Royal Surrey).

Partners

The
ROYAL
MARSDEN
NHS Foundation Trust

NHS
North Central London
Cancer Alliance

IMPERIAL

KING'S
College
LONDON

Supported by

 **uclpartners**
inventing healthcare's future

 **CogStack**


Quantum
Analytica

Introduction

Lung cancer is the leading cause of cancer-related death in the UK, with many cases still diagnosed at a late stage when treatment is less effective. While lung screening is helping detect more cancers earlier, uptake remains lower in some communities, particularly among groups at greater risk of developing the disease.

This Driver Project will use the London Secure Data Environment (SDE) to create a linked, AI-ready dataset that brings together primary care, hospital and screening data. The project will support better understanding of screening uptake, improve early diagnosis and care pathways, and enable the development and validation of AI-enabled risk stratification tools such as LungGuard.

By supporting earlier diagnosis, reducing health inequalities and increasing screening uptake, the project aims to improve outcomes for Londoners while helping the NHS use resources more effectively.

Patient and Public Involvement and Engagement (PPIE)

PPIE is central to this project. A co-applicant with lived experience helped shape the proposal and will continue to support transparency, rigour and public trust. The project team will also work with the Royal Marsden Patient Advisory Group to gather wider patient perspectives.

The project builds on previous engagement with North Central London residents on the use of AI for earlier cancer detection. Participants supported the approach when there was clear patient benefit, while emphasising the importance of transparency, data stewardship, clinical oversight, equity and clear communication. These insights, as well as further patient and public engagement delivered by UCLPartners, will continue to shape the project.

Project Summary – aims and timelines

Aims:

1. Build a linked, analysis-ready lung cancer dataset within the SDE (integrating primary care, screening, and secondary care data).
2. Improve pathway visibility across the lung cancer journey (from GP presentation to diagnosis and treatment).
3. Identify inequalities and drivers of non-attendance in screening.
4. Enable enrichment of structured data using unstructured clinical text (e.g. radiology, pathology, MDT notes via NLP/LLMs).
5. Validate AI risk stratification (AIRS) tools such as LungGuard with real world data and undertake implementation planning including feasibility, adoption pathways, and potential clinical use cases.
6. Develop a scalable governance and technical framework for cross-London data linkage and reuse across other cancer pathways.

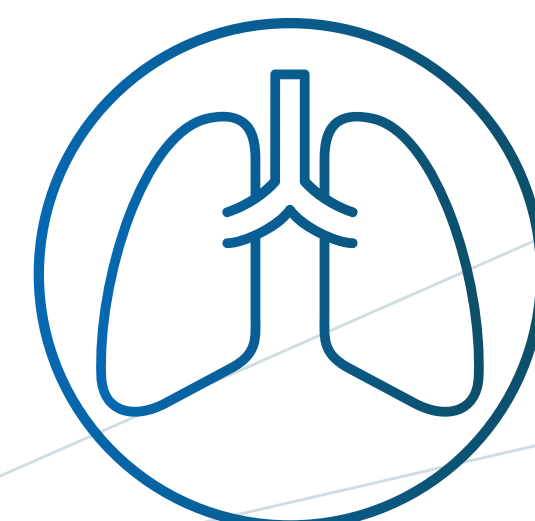
Timeline (over 12 months):

- **0-3:** Define data dictionary, curate primary care dataset, identify screening cohort, data access and retrospective validation for AIRS tools
- **3-6:** Build secondary care datasets, pilot NLP/LLM extraction, progress IG
- **3-12:** Prospective 'shadow-mode' AIRS testing, evaluate real-world feasibility, develop AIRS implementation plan
- **6-9:** Link primary-secondary care data, assess quality, progress DSAs
- **6-12:** Analyse pathways, evaluate clinical utility, deliver governance blueprint

Innovation and desired outcomes

Integrated dataset + validated AIRS model:

- First London-wide linked lung cancer dataset
- Integration of structured + unstructured data
- Real-world validation of AI risk stratification
- Focus on health inequalities / non-attenders



Short-term outcomes:

- Identify and describe high-risk non-attenders
- Improve targeted screening invitations for non-attenders

Medium-term outcomes:

- Define opportunities to improve lung screening clinical data
- Validate AIRS tools in other tumour sites, using SDE data

Long-term outcomes:

- Inform national screening policy
- Scalable framework for other cancers