Project scope: Chest X-ray AI

20 June 2024

Research question

Is artificial intelligence assisted clinical review of chest x-rays effective in improving the detection of lung cancer?

Inclusion criteria

The selection of studies for inclusion in the literature review element of the project will be based on the following criteria:

Population	Patients referred through outpatient services for chest x-ray, with or without suspected lung cancer.
Intervention	Software using AI derived algorithms, alongside clinical review, to find lung abnormalities that may indicate cancer, such as: • Annalise CXR (Annalise ai) • qXR (Qure.ai)
Comparator	A chest X-ray reviewed by a radiologist or radiographer without assistance from AI software.
Outcomes	• Time from chest x-ray to receiving treatment • Time from chest x-ray to referral to CT scan • Time to diagnosis • Number of patients diagnosed with treatable cancers/stage of cancer at detection • Diagnostic accuracy or model performance (sensitivity, specificity, negative predictive value, positive predictive value). • Safety/adverse events • Patients' experiences/views, staff experiences/views, acceptability/ease of use • Costs

Limits	English language Previous search conducted between 3rd and 4th January 2024, excluded studies published prior to 2004 Update to be run from 5th January 2024 to current date.

Equity

Incidence rates for lung cancer in the UK are highest in people aged 85 to 89 (Cancer Research UK 2016-2018). Lung cancer occurs more frequently in men than in women. Rates are disproportionately increasing in women. In women lung cancer is more common in white women and in men lung cancer is more common in white men and Bangladeshi men. Lung cancer is more common in deprived groups and in groups that may not engage as often with health services.

AI algorithms developed exclusively (or largely) on one group may generalize poorly beyond that population leading to bias in their application. For example, a study used two large x-ray datasets to vary the level to which women were underrepresented in an AI model diagnosing thoracic diseases, this led to lower accuracy rates in women, until sufficient women were added to the underlying model data set. Another study demonstrated the importance of a socioeconomically and ethnically diverse underlying training set for an AI model that examined patients x-rays to measure the severity of osteoarthritis and associated pain. The recently published Equity in Medical Devices: independent review recommends that manufacturers report on the diversity of data used to train the AI algorithm and are transparent about any limitations in device performance across different groups. Focused research may be required to explore any such limitations.

High quality X-ray may be difficult to obtain in some groups e.g. people with morbid obesity or scoliosis. Lower quality images may be rejected by AI software because it is unable to interpret them, this will depend on how the AI algorithms have been trained and may disadvantage some groups (NICE, 2023). Lung abnormalities may be more difficult to identify in certain groups, including younger women who do not smoke, people with asthma or chronic obstructive pulmonary disease (COPD), and people 'whose family background means they may be at higher risk of having lung cancer' (NICE, 2023). If AI supported clinical review of chest x-rays aids in the detection of lung cancer, this may be especially beneficial to these groups.

Planned activities

SHTG have agreed on the following activities to support the development of an SHTG Assessment on artificial intelligence software to assist in the analysis of Chest X-rays to identify suspected lung cancer:

1. Evidence review of the published literature on the clinical effectiveness, cost effectiveness, safety and patient aspects

2. An evaluation of annalise.ai based on 12 months prospective real-world data from NHS Grampian.

3. A summary of a report provided to SHTG from the RADICAL study being led by NHS Greater Glasgow and Clyde (including any interim data analyses provided by the RADICAL trial investigators).

3. A cost-effectiveness analysis of annalise.ai versus the comparator using real-world data from NHS Grampian.

4. A plain language version.

5. Engagement with clinical experts during peer review.

End products

At the end of the project, SHTG will publish:

- SHTG Assessment
- An evaluation report as supplementary material
- Plain language summary
- Expert comments from peer review

Timescales (approximate)

Approximate publication date November 2024.