Early Value Assessment report commissioned by the NIHR Evidence Synthesis Programme on behalf of the National Institute for Health and Care Excellence

Title of project: Artificial intelligence software for analysing chest X-ray images to identify suspected lung cancer

Produced by: Warwick Evidence

Lead Authors: Jill Colquitt/ Mary Jordan
Co-authors: Rachel Court, Emma Loveman, Janette Parr, Iman Ghosh, Peter Auguste, Mubarak Patel, Chris Stinton

Correspondence to: Dr Chris Stinton
Division of Health Sciences
Warwick Medical School
University of Warwick
Coventry CV4 7AL

Tel: 02476 574 701
Email: c.stinton@warwick.ac.uk

Date completed: 17.02.2023

This EVA was conducted against a background of industrial action in universities across the UK. This reduced the time available for this project.

Source of funding:

This report was commissioned by the NIHR Evidence Synthesis Programme as project number NIHR135755.

Declared competing interests of the authors

None of the authors have any competing interests.

Acknowledgements

The authors would like to thank NICE Specialist Committee Members for their expert clinical advice, the Radiology Department, Respiratory Medicine Department and the Lung Cancer

Team at the Royal United Hospitals Bath NHS Foundation Trust for their data provision and advice on departmental practices, and Dr Dan Todkill for comments on the report.

Ethics statement:

This EVA consists of secondary research, therefore ethical approval was not required.

Information Governance Statement

University of Warwick is committed to handling all personal information in line with the UK Data Protection Act (2018) and the General Data Protection Regulation (UK GDPR).

Under the Data Protection legislation, the University of Warwick is the Data Controller, and you can find out more about how we handle personal data, including how to exercise your individual rights and the contact details for our Data Protection Officer here https://warwick.ac.uk/services/legalandcomplianceservices/dataprotection/

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Artificial intelligence software for analysing chest X-ray images to identify suspected lung cancer. Early Value Assessment commissioned by the NIHR Evidence Synthesis Programme on behalf of the National Institute for Health and Care Excellence.

Contributions of authors

Rachel Court (Information specialist) developed the search strategy and undertook searches.

Jill Colquitt (Honorary Research Fellow), Emma Loveman (Honorary Research Fellow), Iman Ghosh (Research Associate), and Janette Parr (Research Associate) conducted the clinical effectiveness review.

Mary Jordan (Research Fellow) and Peter Auguste (Assistant Professor) contributed to the cost-effectiveness review and developed the conceptual model.

Mubarak Patel (Research Associate) developed the analysis plan.

Chris Stinton (Senior Research Fellow) led the project.

All authors were involved in developing the protocol and writing the draft and final version of the report.

Academic and commercial in confidence information:

Please note that throughout the report academic in confidence (AIC) information is yellow and underlined and commercial in confidence (CIC) information is marked blue and underlined.

Data sharing statement:

All available data can be obtained by contacting the corresponding author.

Word count: 21,431

Abstract

Background

The objectives of this early value assessment (EVA) were to identify evidence on adjunct AI software for analysing chest x-rays from people referred from primary care for suspected lung cancer, identify evidence gaps to help direct data collection and further research, and develop a conceptual model to inform discussion of what would be required to develop a fully-executable cost-effectiveness model for future economic evaluation.

Objectives

We conducted searches for evidence in nine electronic databases and through consultation with clinical experts. The review methods followed rapid evidence synthesis methods.

Methods

We conducted searches for evidence in nine electronic databases and through consultation with clinical experts. The review methods followed rapid evidence synthesis methods.

Results

We did not find any evidence that met the inclusion criteria for this EVA.

Conclusions

There is no applicable evidence on which to evaluate the impact of adjunct AI software for analysing chest x-rays from people referred from primary care for suspected lung cancer. We provide examples of research recommendations.

Plain English summary

Lung cancer is one of the most common types of cancer in the UK. Early diagnosis of lung cancer may improve survival, but lung cancer is often diagnosed late. Chest X-rays can be used to look for features of lung cancer in people with signs or symptoms of lung cancer. In some cases, features of lung cancer are found in chest X-rays taken for other reasons. There can be delays in getting an X-ray because of high demand and there can also be delays in the X-ray being read. Sometimes features of lung cancer are missed on X-ray.

Artificial intelligence is a part of computer science where computers do some tasks that are usually done by humans. Artificial intelligence software may help by finding features of lung cancer in a chest X-ray. A radiologist can then look at these X-rays. There are different types of artificial intelligence software available but a lack of information about how lung cancer diagnosis could change if they are used. There is also a lack of information about the costs of these software to the NHS.

This project looked at the evidence on the use of artificial intelligence software in the diagnosis of lung cancer for people referred from primary care. There were no studies that looked at this topic among people from primary care. We summarised the closest evidence that we could find instead. All of this had flaws in how the research was designed or conducted. This meant we could not tell if the results were accurate or how helpful they were to the topic of this review. It was not clear if artificial intelligence helped to find cancers. There was no evidence on if artificial intelligence could improve people's health.

We made a theoretical model so we could discuss the best way to assess if artificial intelligence software might be cost-effective in detecting lung cancer and what evidence would be needed to be able to do this in a fully working model. We found no evidence to link artificial intelligence software with outcomes after diagnosis, e.g. people's health.

We worked out the cost of adding artificial intelligence software to review chest X-rays in people referred from their general practitioner, for the first five years, based on one NHS Trust.

We found no evidence that we could include that answered the question of whether artificial intelligence software could help radiologists review chest X-rays of people referred from a general practitioner. We have made recommendations on future studies that will be important to answer this question.

Scientific summary

Background

Lung cancer occurs when abnormal cells multiply in an uncontrolled way to form a tumour in the lung. It is one of the most common types of cancer in the UK and each year over 43,000 new cases are diagnosed. In the early stages of the disease people usually do not have symptoms which means lung cancer is often diagnosed late. The five-year survival rate for lung cancer is low, at below 10%. Early diagnosis may improve survival. NICE has identified software that has an artificial intelligence (AI) developed algorithm (referred to hereafter as AI software) as potentially useful in assisting with the identification of suspected lung cancer. AI combines computer science and datasets to enable problem solving. Machine learning and deep learning are sub-fields of AI. They comprise AI algorithms which seek to create expert systems to make predictions or classifications based on data input.

This assessment covers the use of AI software as an adjunct to an appropriate radiology specialist (radiologist, reporting radiographer) to assist in the identification of suspected lung cancer on chest x-rays (CXR). AI technologies subject to this assessment are standalone software platforms developed with deep learning algorithms to interpret CXR. The algorithms are fixed but updated periodically. The AI software automatically interprets radiology images from the CXR to identify abnormalities or suspected abnormalities. The abnormalities detected and the methods of flagging the location and type of abnormalities differ between different AI technologies. For example, a CXR may be flagged as suspected lung cancer when a lung nodule, lung mass, hilar enlargement, or a combination of these are identified. A technology may classify CXRs into those with and without a nodule, or it may identify several different abnormalities or lung diseases.

Objectives

The overall aim of this early value assessment (EVA) is to identify evidence on adjunct AI software for analysing CXR for suspected lung cancer, and identify evidence gaps to help direct data collection and further research. A conceptual modelling process will be undertaken to inform discussion of what would be required to develop a fully-executable cost-effectiveness model for future economic evaluation. The assessment is not intended to

replace the need for a full assessment (Diagnostic Assessment Report) or to provide sufficient detail or synthesis to enable a recommendation to be made on whether AI software can be implemented in clinical practice at the present time.

There are two populations of interest in this EVA: (1) people referred from primary care who are having CXR because they have symptoms suggestive of lung cancer (symptomatic population), and (2) people referred from primary care who are having CXR for reasons unrelated to lung cancer (incidental population). Based on the scope produced by NICE we defined the following questions to inform future assessment on the benefits, harms, and costs of adjunct AI for analysing on CXR for suspected lung cancer compared to human reader alone in these populations:

1. What is the test accuracy and test failure rate of adjunct AI software to detect lung cancer on CXR?
2. What are the practical implications of adjunct AI to detect lung cancer on CXR?
3. What is the clinical effectiveness of adjunct AI software applied to CXR?
4. What are the cost and resource use considerations relating to use of adjunct AI to detect lung cancer?
5. What would a health economic model to estimate the cost-effectiveness of adjunct AI to detect lung cancer look like?

Methods

Data sources

MEDLINE All (via Ovid), Embase (Ovid), Cochrane Database of Systematic Reviews (Wiley), Cochrane CENTRAL (Wiley), Epistemonikos, ACM Digital Library, WHO ICTRP, clinical experts.

Eligibility criteria

Population: people referred for CXR from primary care because they have symptoms suggestive of lung cancer, people referred for CXR from primary care for reasons unrelated to lung cancer

Intervention: radiology specialist (radiologist, reporting radiographer) with adjunct AI

Comparator: radiology specialist (radiologist, reporting radiographer) without adjunct AI

Outcomes: test accuracy, patient management, clinical effectiveness

Study selection, data extraction, assessment of risks of bias

Titles and abstracts of all identified records were screened by one reviewer against the review eligibility criteria, with a random 20% screened by a second reviewer. Full texts of records considered potentially relevant by either reviewer were retrieved and assessed for inclusion by one reviewer. A random 20% sample were assessed independently by a second reviewer, with any disagreements resolved by consensus or discussion with a third reviewer. We planned to extract data into a piloted form, assess risk of bias, and synthesise data using methods described in the research protocol; however, no studies met the inclusion criteria. Post hoc methods were determined following discussions with NICE to select and summarise the closest available evidence to the review inclusion criteria. Studies that assessed eligible AI software in conjunction with radiology specialist versus radiology specialist alone; but where the referral status and symptomatic status of the population was unclear, were selected. Data were extracted by one reviewer, with a random 20% checked by a second reviewer. Results were summarised narratively, and key biases were noted.

Data synthesis

Narrative data synthesis was performed.

Modelling

The conceptual modelling process explored both the structure, and evidence requirements for parameter inputs, for future model development. An iterative approach was taken to facilitate identification of cost outcomes, potential value drivers for AI software for this indication, and evidence linkage requirements for longer term outcomes, where time allowed. Costs associated with implementing AI software were also considered.

Information to inform the conceptual model were obtained from a variety of sources including literature review, current clinical guidelines, discussion with specialist clinical experts, and the companies submitting evidence submissions on AI software.

Given the time available, the diagnostic component of the model was the primary focus of the health economics aspect of the report. Priority was given to the following considerations:

- Input parameters to populate model – including consideration of type of evidence required, sources available, and gaps in the evidence.
- Relevant outcome measures to compare cost and clinical effectiveness of AI software in the detection of lung cancer.
- Identification of potential value drivers of model – with recommendations of how these can be measured for inclusion in a cost-effectiveness model.

Results

Test accuracy, practical implications, and clinical effectiveness

No studies met the inclusion criteria of the review. Two ongoing studies with unclear eligibility were identified. In the absence of available evidence, we summarised data from six studies that had unclear populations but included a comparison of CXR read by readers with and without the use of commercial AI software.

Statistical comparisons were not undertaken in most of the studies, but there was some evidence that sensitivity might be higher amongst specialist radiologist with AI than specialist radiologist without AI. This finding was not consistent between studies, however. No significant differences were observed for specificity, positive predictive value, or number of cancers detected. None of the studies provided evidence on the clinical effectiveness of adjunct AI software. The summarised excluded studies were small retrospective studies with important methodological limitations and their generalisability to the UK population is unclear.

Conclusions

There is currently no evidence on the use of adjunct AI software for the detection of suspected lung cancer on CXR in either people referred from primary care with symptoms of lung cancer or in people referred from primary care for other reasons.

Implications for service provision

Lung cancer pathways are complex and contain many routes to diagnosis. Whilst national guidance and timelines for diagnosis exist, practice variation is widespread throughout

radiology departments and lung cancer teams both within and across NHS Trusts. With many ways to achieve these targets, changes in any area of the diagnostic pathway may have significant impact elsewhere.

There is some evidence about the impact of CXR results on the diagnostic pathway when performed without AI assistance, as is current practice. This is limited and difficult to compare results due to the different study designs used and different outcomes reported. There is no published evidence to link measures of progression through the diagnostic pathways with long-term outcomes such as stage at diagnosis and survival.

There is currently no evidence to show the impact that the addition of AI software to CXR review has on the diagnosis of lung cancer. There may be multiple ways AI software could change measures along this pathway. These could include improved accuracy of lung cancer detection directing patients along the quickest pathway to diagnosis, quicker report turnaround time to achieve earlier confirmatory testing, or prioritisation of cases for review including those without lung cancer who can be discharged more quickly and free up staff time and resources. AI software may also impact pathways negatively by increasing the number of lung nodules detected which are benign, increasing the number of patients undergoing a CT scan which they may not have needed. This would be detrimental to the patient with increased exposure to radiation and anxiety due to a positive CXR result, and have cost and resource use implications affecting the department.

With a lack of evidence on AI software, the impact on service provision is unknown and may have significant implications in terms of progression through diagnostic pathways, resource use, costs, and patient outcomes.

Contents

List of Appendices

List of figures

List of tables

List of abbreviations

AI	Artificial intelligence
ASG	Assessment subgroup
CXR	Chest x-ray
DAC	Diagnostic Advisory Committee
EAG	External Assessment Group
EVA	Early value assessment
NHS	National Health Service
NICE	The National Institute for Health and Care Excellence
UK	United Kingdom

Definitions of terms

Term	Definition
Artificial Intelligence (AI)	The ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings.
CADe	Computer-aided detection
CADx	Computer-aided diagnosis
CAST	Computer-aided simple triage
Deep learning	Deep learning is a method in artificial intelligence (AI) that teaches computers to process data in a way that is inspired by the human brain. Deep learning models can recognize complex patterns in pictures, text, sounds, and other data to produce accurate insights and predictions.
False negative value	The number of cases in which the index test has wrongly suggested the patient as being disease-free when they do have the disease. FN= c
False positive value	The number of cases in which the index test has wrongly indicated the patient as having the disease when they do not have the disease. FP= b
Ground Truth	Ground truth refers to the actual nature of the problem that is the target of a machine learning model, reflected by the relevant data sets associated with the use case in question.
Machine Learning	In artificial intelligence (a subject within computer science), discipline concerned with the implementation of computer software that can learn autonomously.
Reference standard	The test, combination of tests, or procedure that is considered the best available method of categorising participants in a study of diagnostic test accuracy as having or not having a target condition.
Sensitivity	The proportion of people who test positive for a disease amongst people who have the disease of interest. The ratio between the true positive value and (true positive value + false negative value). $$Sensitivity = \frac{a}{a+c} = \frac{TP}{TP+FN}$$
Specificity	The proportion of people who test negative for a disease amongst people who do not have the disease of interest. The ratio between the true negative value and (true negative value + false positive value).

	$Specificity = \frac{d}{b+d} = \frac{TN}{TN+FP}$
Survival rate	The percentage of people in a study or treatment group who are still alive for a certain period of time after they were diagnosed with or started treatment for a disease, such as cancer.
True negative value	The number of cases in which the index test has correctly indicated the patient as being disease-free. TN= d
True positive value	The number of cases in which the index test has correctly indicated the patient as having the disease. TP= a

1. Introduction

1.1. Purpose of the decision to be made

Lung cancer occurs when abnormal cells multiply in an uncontrolled way to form a tumour in the lung.[1] It is one of the most common types of cancer in the UK and each year over 43,000 new cases are diagnosed.[2] In the early stages of the disease people usually do not have symptoms which means lung cancer is often diagnosed late.[3] The five-year survival rate for lung cancer is low, at below 10%.[2] Early diagnosis may improve survival.[3] NICE has identified software that has an artificial intelligence (AI) developed algorithm (referred to hereafter as AI software) as potentially useful in assisting with the identification of suspected lung cancer.

The purpose of this early value assessment (EVA) is to assess the evidence on adjunct AI software for analysing chest x-rays (CXR) for suspected lung cancer, and identify evidence gaps to help direct data collection and further research. A conceptual modelling process will be undertaken to inform discussion of what would be required to develop a full-executable cost-effectiveness model for future economic evaluation.

1.2. Population

There are two populations of interest in this EVA: (1) people referred from primary care who are having CXR because they have symptoms suggestive of lung cancer (symptomatic population), and (2) people referred from primary care who are having CXR for reasons unrelated to lung cancer (incidental population).

1.3. Condition

Lung cancer is one of the most common causes of cancer in the UK. There approximately 43,000 new cases diagnosed annually.[2] The incidence of lung cancer is highest amongst older people.[4] It is rare in people under the age of 40. More than 40% of people diagnosed with lung cancer are 75 years or older.[3]

Lung cancer occurs when abnormal cells multiply in an uncontrolled way to form a tumour in the lung.[1] Cancer that begins in the lungs is called primary lung cancer. Cancer that begins elsewhere and spreads to the lungs is called secondary lung cancer. There are two main forms of primary lung cancer: non-small-cell lung cancer and small-cell lung cancer. These are named after the type of cell in which the cancer started growing. Non-small-cell lung

cancer is the most common type (80-85% of cases) and can be classified into one of three kinds: squamous cell carcinoma, adenocarcinoma, or large-cell carcinoma. Small-cell lung cancer is less common but usually spreads faster than non-small-cell lung cancer.[3] Most cases of lung cancer are caused by smoking. Although people who have never smoked can also develop the condition, smoking cigarettes is responsible for more than 70% of cases.[3] People who smoke are 25 times more likely to get lung cancer than people who do not smoke. Other exposures can also increase the risk of lung cancer. These include radon gas (naturally occurring), occupational exposure to certain chemicals and substances, and pollution.[3]

Symptoms of lung cancer include persistent cough, coughing up blood, and shortness of breath. However, in the early stages of the disease people usually do not have symptoms.[3] This means lung cancer is often diagnosed late. In 2018, more than 65% of lung cancers in England were diagnosed at stage 3. Survival rates for lung cancer are very low. Recent estimates suggest 5-year survival rates of 10%.[3] The NHS Long Term Plan sets out the NHS's ambition to diagnose 75% of all cancers at an early stage by 2028.[5]

1.4. Technologies under assessment

AI combines computer science and datasets to enable problem solving. Machine learning and deep learning are sub-fields of AI. They comprise AI algorithms which seek to create expert systems to make predictions or classifications based on data input.[6] Many paradigms of deep learning have been developed but the most used of these is the Convolutional Neural Network.[7]

This assessment covers the use of AI software as an adjunct to an appropriate radiology specialist (radiologist, reporting radiographer) to assist in the identification of suspected lung cancer. AI technologies subject to this assessment are standalone software platforms developed with deep learning algorithms to interpret CXR. The algorithms are fixed but updated periodically. The AI software automatically interprets radiology images from the CXR to identify abnormalities or suspected abnormalities. The abnormalities detected and the methods of flagging the location and type of abnormalities differ between different AI technologies. For example, a CXR may be flagged as suspected lung cancer when a lung nodule, lung mass, hilar enlargement, or a combination of these are identified. A technology

may classify CXRs into those with and without a nodule, or it may identify several different abnormalities or lung diseases**.**

1.5. Comparators

The comparator for this assessment is CXR images reviewed by an appropriate radiology specialist (radiologist, reporting radiographer) without assistance from AI software.

1.6. Reference standards

Following CXR, people with suspected lung cancer should be offered a contrast-enhanced chest CT scan to diagnosis and stage the disease (contrast medium should only be given with caution to people with known renal impairment). The liver, adrenals and lower neck should also be included in the scan.[8] If the CT scan indicates there may be cancer, the type and sequence of investigations may vary but typically include a positron emission tomography-CT (PET-CT) scan and an image-guided biopsy. Other methods that may be used include magnetic resonance imaging (MRI), endobronchial ultrasound-guided transbronchial needle aspiration (EBUS-TBNA), and endoscopic ultrasound-guided fine-needle aspiration (EUS-FNA).[8] The PET-CT scan can show where there are active cancer cells which can help with diagnosis and choosing the best treatment.[3]

1.7. Care pathway

The identification of people with signs and symptoms suggestive of lung cancer often happens in primary care. The NICE guideline on recognition and referral for suspected lung cancer recommends that people aged 40 and over are offered an urgent CXR (within 2 weeks of referral) if they have two or more symptoms of lung cancer, or if they have ever smoked and have at least one of the following unexplained symptoms: cough, fatigue, shortness of breath, chest pain, weight loss, appetite loss.[9] An urgent CXR should also be considered for people aged 40 or over if they have persistent or recurrent chest infection, finger clubbing, enlarged lymph nodes near the collarbone or in the neck (supraclavicular lymphadenopathy or persistent cervical lymphadenopathy), chest signs consistent with lung cancer, or increased platelet count (thrombocytosis). If the CXR findings suggest lung cancer, referral to secondary care should be made using a suspected cancer pathway referral for an appointment within 2 weeks. If the CXR is normal (without any clinically relevant lung abnormalities), high risk patients, i.e., those who present with ongoing, unexplained

symptoms, are referred to secondary care. Low risk patients are discharged. In this EVA, AI software is applied to CXR of patients who are referred from primary care. Referrals for CXR outside primary care are beyond the scope of this project.

2. Decision questions and objectives

The overall aim of this project was to identify evidence on adjunct AI software for analysing CXR for suspected lung cancer, and identify evidence gaps to help direct data collection and further research. A conceptual modelling process was undertaken to inform discussion of what would be required to develop a full-executable cost-effectiveness model for future economic evaluation. These were examined via an EVA. The assessment was not intended to replace the need for a full assessment (Diagnostic Assessment Report) or to provide sufficient detail or synthesis to enable a recommendation to be made on whether AI software can be implemented in clinical practice at the present time.

Based on the scope produced by NICE[10] we defined the following questions to inform future assessment on the benefits, harms, and costs of adjunct AI for analysing on CXR for suspected lung cancer compared to human reader alone:

1. What is the test accuracy and test failure rate of adjunct AI software to detect lung cancer on CXR?
2. What are the practical implications of adjunct AI to detect lung cancer on CXR?
3. What is the clinical effectiveness of adjunct AI software applied to CXR?
4. What would a health economic model to estimate the cost-effectiveness of adjunct AI to detect lung cancer look like?
5. What are the cost and resource use considerations relating to use of adjunct AI to detect lung cancer?

3. Methods

The review is registered on PROSPERO (registration number CRD42023384164), and the protocol is available from the NICE website (https://www.nice.org.uk/guidance/gid-dg10065/documents/final-protocol).

The timeline to produce this EVA report was ten weeks, which is substantially shorter than a typical systematic review or rapid review. To achieve the aims within the timeline, pragmatic decisions regarding the methods were made in collaboration with NICE and clinical experts.

3.1. Methods for assessing test accuracy, practical implications, and clinical effectiveness

3.1.1. Search strategy

An iterative approach was taken to develop the search strategy, making use of relevant records identified during initial scoping searches and from relevant reviews.[11, 12] The strategy was developed by an information specialist, with input from team members, aiming for a reasonable balance of sensitivity and specificity. Based on scoping work already undertaken, a series of complementary, targeted searches were favoured over a single search to retrieve a manageable number of records to screen (see Appendix 1). Searches were run in a range of relevant bibliographic databases covering the fields of medicine and computer science. Searches were limited to studies published in English because studies published in other languages are likely to be difficult to assess in the timescale of this EVA. Non-human studies, letters, editorials, communications, and conference abstracts were removed during the searches. No date limit was applied to the searches, but only records published in or after 2012 were screened. Database search strings were developed for MEDLINE and appropriately translated for each of the other databases, considering differences in thesaurus terms and syntax. The following bibliographic databases were searched: MEDLINE All (via Ovid), Embase (Ovid), Cochrane Database of Systematic Reviews (Wiley), Cochrane CENTRAL (Wiley), Epistemonikos, ACM Digital Library.

A search for ongoing trials were conducted in the WHO ICTRP. A search for ongoing systematic reviews was undertaken in the PROSPERO database.

The full record of searches is provided in Appendix 1.

Records were exported into EndNote X9.3, where duplicates were systematically identified and removed. Reference lists of included studies and a selection of relevant reviews were

checked. Experts and team members were consulted and encouraged to share relevant studies.

3.1.2. Eligibility criteria

The eligibility criteria for the test accuracy, practical implications, and clinical effectiveness questions are presented in Table 1.

Table 1: Eligibility criteria

	Key question 1. What are the test accuracy and test failure rates of adjunct AI software to detect lung cancer on CXR? Sub-questions: 1a. What is the test accuracy of adjunct AI software to detect lung nodules? 1b. What is the concordance in lung nodule detection between radiology specialist with and without adjunct AI software	Key question 2. What are the practical implications of adjunct AI software to detect lung cancer on CXR?[a]	Key question 3. What is the clinical effectiveness of adjunct AI software applied to CXR?
Population	Adults referred from primary care who are: 1. undergoing CXR due to symptoms suggestive of lung cancer, e.g., cough, fatigue, shortness of breath, chest pain, weight loss, appetite loss, persistent or recurrent chest infection, finger clubbing, supraclavicular lymphadenopathy or persistent cervical lymphadenopathy, chest signs consistent with lung cancer and/or thrombocytosis (symptomatic population) 2. undergoing CXR for reasons unrelated to lung cancer (incidental population) Where data permits, subgroups will be considered based on: • Ethnicity • Age • Sex • Socio-economic status		
Target condition	Lung cancer		
Intervention	CXR interpreted by radiology specialist (radiologist, reporting radiographer) in conjunction with the following AI software: AI-Rad Companion Chest X-ray (Siemens Healthineers), Annalise CXR (annalise.ai), Auto Lung Nodule Detection (Samsung), ChestLink Radiology Automation (Oxipit), ChestView (GLEAMER), Chest X-ray (Rayscape), ClearRead Xray – Detect (Riverain Technologies), InferRead DR Chest (Infervision), Lunit INSIGHT CXR (Lunit), Milvue Suite (Milvue), qXR (Qure.ai), red dot (behold.ai), SenseCare-Chest DR Pro (SenseTime), VUNO Med-Chest X-Ray (VUNO)		
Comparator	CXR interpreted by radiology specialist (radiologist, reporting radiographer) without the use of AI software		
Reference standard	For accuracy of lung cancer detection: Lung cancer confirmed by histological analysis of lung biopsy, or diagnostic methods specified in NICE guideline 122, [8] where biopsy is not applicable	NA	NA

	For accuracy of nodule detection: Radiology specialist (single reader or consensus of more than one reader)		
Outcome	Test accuracy for the detection of lung cancer (sensitivity, specificity, positive predictive value, numbers of true positive, false positive, true negative, false negative results, number of lung cancers diagnosed) Test failures (rates, and data on inconclusive, indeterminate, and excluded samples, failure due to any other reason) Characteristics of discordant cancers cases Test accuracy for the detection of lung nodules Concordance in lung nodule detection between radiology specialist with and without adjunct AI software	Practical implications[a] (time to x-ray report, CT scan, diagnosis, turnaround time (image review to radiology report), acceptability of software to clinicians, impact on clinical decision-making, impact of false positives on workflow)	Mortality, morbidity, health-related quality of life
Study design	Comparative study designs		
Publication type	Peer reviewed papers		
Language	English		
Exclusion	Versions of AI software that are not commercially available, are not named in the protocol, or are not specified in the study publication. Computer aided detection that does not include AI software. Non-human studies. Letters, editorials, communications, conference abstracts, qualitative studies. People with a known diagnosis of lung cancer at the time of CXR. Studies of children. Study designs that do not include a control/comparator arm. Simulation studies or studies using synthetic images. Studies not applicable to primary care patients, e.g., neurosurgery, transplant, or plastic surgery patients, people in secure forensic mental health services. Studies where more than 10% of the sample do not meet our inclusion criteria. Studies without extractable numerical data. Studies that provided insufficient information for assessment of methodological quality/risk of bias. Articles not available in the English language. Studies using index tests or reference standards other than those specified in the inclusion criteria. Studies of people who do not have signs and symptoms of cancer or a suspected condition or trauma (i.e., people undergoing health screening). Studies where it cannot be determined if the inclusion criteria are met.		

[a] For the 'acceptability' and 'impact on decision-making' outcomes, the relevant population is the radiologist or reporting radiographer interpreting the CXR of adults defined under 'Population'.

3.1.3. Review strategy

Titles and abstracts of records identified by the searches were screened by one reviewer, with a random 20% assessed independently by a second reviewer. Records considered potentially relevant by either reviewer were retrieved for further assessment. Full text articles were assessed against the full inclusion/exclusion criteria by one reviewer. A random 20% sample were assessed independently by a second reviewer. Disagreements were resolved by consensus, or through discussion with a third reviewer. Records rejected at full text stage (including reasons for exclusion) are report in Appendix 2.

3.1.4. Data extraction

We planned to extract data into a piloted electronic data collection form. Data were to be extracted by one reviewer, with a random 20% checked by a second reviewer, and disagreements resolved by consensus or discussion with a third reviewer. However, no studies met the inclusion criteria.

3.1.5. Risk of bias

We planned to assess risk of bias of included studies using tools appropriate to the study design, such as those produced by the Joanna Briggs Institute (JBI).[13]. Risk of bias was to be assessed by one reviewer, with a random 20% assessed by a second reviewer and disagreements resolved through consensus or discussion with a third reviewer. As no studies met the inclusion criteria, no formal risk of bias assessment was undertaken.

3.1.6. Analysis and synthesis

Methods of analysis and synthesis were described *a priori* in the research protocol.[14] However, no studies met the inclusion criteria, so no data synthesis was undertaken.

3.1.7. Post hoc methods

No studies meeting the inclusion criteria were identified. Following discussions with the NICE Technical team for this project, we examined the list of excluded studies that were closest to the review inclusion criteria (see Table 1), i.e.:

- Interventions: CXR interpreted by radiology specialist (radiologist, reporting radiographer) in conjunction with eligible AI software versus radiologists alone and/or reference standard

- Population: no details provided on the referral status or symptom status (studies that had an explicitly excluded population, e.g. health screening, pre-operative CXR, inpatients, Accident & Emergency were not selected)
- Outcomes: as defined in Table 1

Selected studies were tabulated using the approach described in section 3.1.4 and key biases were noted. Results were summarised narratively.

3.2. Methods for developing a conceptual cost-effectiveness model

This section describes the process, methods and rationale for the development of a conceptual[15] decision analytic model to inform potential full cost-effectiveness evaluation of adjunct AI software for analysing CXR images to identify suspected lung cancer.

The conceptual modelling process explored both the structure, and evidence requirements for parameter inputs, for future model development. This was to facilitate identification of cost outcomes, potential value drivers for AI software for this indication, and evidence linkage requirements for longer term outcomes. Costs associated with implementing AI software were also considered.

Information to inform the conceptual model were obtained from a variety of sources including literature review, current clinical guidelines, discussion with specialist clinical experts, and the companies submitting AI software for assessment.

3.2.1. Literature review

A pragmatic search of the literature was used to identify existing methods of cost-effectiveness modelling for AI software in CXR and inform parameterisation of the conceptual model. It was not intended as a substitute for a systematic literature review, nor to provide a definitive summary of evidence gaps. This will be required for any future development of an executable cost effectiveness model.

Following initial scoping searches, we did not expect to find any full economic evaluations of AI software as an adjunct to radiology specialist review of CXR, particularly in the primary care population. For this reason, a broad search strategy was used across two databases (Medline and Tufts CEA), and broad screening criteria applied. The primary inclusion criterion was “lung cancer studies”, but following this any study which could inform the

structure or parameters of a conceptual model were identified at title/abstract level. Full text assessment of these papers was used to refine screening criteria further into studies which satisfied (1) the primary care referral population, (2) those with specific intention of diagnosis or screening, and (3) those most relevant to the UK setting. Reference lists of these studies and publication lists of authorship groups were also screened for any further potentially relevant papers. Studies identified in these targeted reviews were not subject to a formal assessment but discussed narratively. This focused on the methods used, assumptions made, availability of evidence to support evidence-linkage approaches, and considerations for future modelling and research.

3.2.2. Clinical guidelines

The structure of the decision analytical model is intrinsically linked to current clinical pathways. Key points throughout the clinical pathway for detection and management of lung cancer, and the positioning of AI software within this pathway (for adults referred for CXR from primary care), were identified with reference to Figure 1 in the final NICE scope for this topic,[10] existing guidelines on the diagnostic and care pathway,[8, 9, 16, 17] and close collaboration with clinical experts.

3.2.3. Clinical expert involvement

Information on the relevant AI technologies under review were obtained from company submissions, with requests for additional information sent to companies that registered as stakeholders (Annalise AI, Behold AI, Infervision, Lunit Inc. and Siemens Healthcare).

Using the information gathered from these sources, an iterative process was used to achieve a model structure that is pragmatic in its representation of the complex clinical pathways that adults from primary care populations may follow to arrive at a diagnosis of lung cancer.

Given the time available to conduct this EVA, the primary focus of this report was on the diagnostic component of the model. Priority was given to the following:

- Input parameters to populate model – including consideration of the type of evidence required, sources available, and gaps in the evidence.
- Relevant outcome measures to compare cost and clinical effectiveness of AI software in the detection of lung cancer.

- Identification of potential value drivers of model – with recommendations of how these can be measured for inclusion in a cost-effectiveness model.

Once diagnosis is achieved in the model, evidence linkage between intermediate outcomes and long-term outcomes is required to assess cost-effectiveness over a clinically appropriate time horizon. These mainly relate to the mapping of the disease state (i.e., lung cancer), and are not specific to the diagnostic technology being assessed (e.g., utilities, costs, and effects of current treatments). Potential sources for the main longer-term outcomes were identified during the literature search, with focus on those relevant to the UK setting and in line with requirements of the NICE reference case.[18] An overview is presented in this report as an example of current practices in modelling lung cancer.

3.2.4. Methods to assess potential budget impact

Estimates of the potential budget impact of introducing AI software as an adjunct to radiology specialist review of CXR were calculated based on methods for a budget impact analysis (BIA) outlined in the NICE evidence standards framework for digital health technologies[19] and ISPOR Task Force recommendations.[20] These identify six key elements which require inputs for the modelling framework of a BIA:

- Size and characteristics of affected population;
- Current intervention mix without the new intervention;
- Costs of current intervention mix;
- New intervention mix with the new intervention;
- Cost of the new intervention mix;
- Use and cost of other health conditions, and treatment-related health-care services[20]

Given the limitation of time and scope, a fully comprehensive BIA was not attempted as this would have required data on any changes in resource use and associated cost. The intended outcome of this report was a conceptual model where no outcome data was run, or results produced. Therefore, estimates on this element were not included. The aim was to approximate the budget impact at an individual institution level, with information sourced from the literature and supplemental information provided by representatives of the institution used as an example.

Company submissions to NICE as part of the Diagnostic Assessment Programme request for

information were screened for cost data. Clarifying questions were sent to all companies (whether costings were already submitted or not) to obtain more granular detail for the purpose of budget impact analysis.

Records retrieved in the broad cost search were screened at title/abstract level by one reviewer (MJ) to identify any studies which may have been applicable. These were then retrieved as full text and their suitability for use assessed. Studies which yielded relevant information were retained, data extracted, and authors contacted to obtain further context specific information.

4. Results: test accuracy, practical implications, and clinical effectiveness

4.1. Results of literature searches

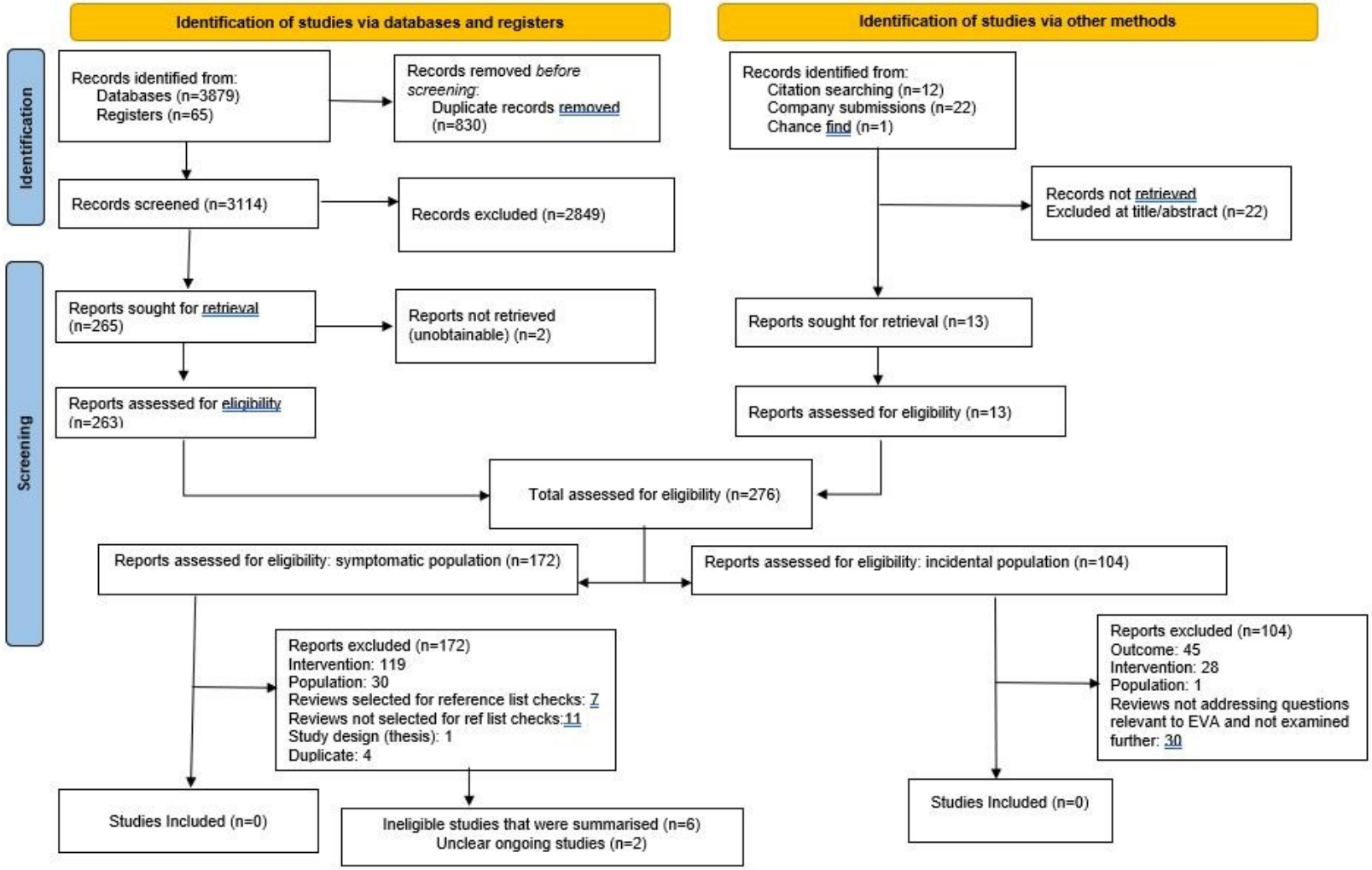
Figure 1 shows the flow of studies through this review. Searches identified a total of 3149 records. Of these, 172 were identified as potentially relevant to the symptomatic population and 104 to the incidental population. Full texts were obtained and screened. None of the studies met the inclusion criteria specified in Table 1. The eligibility of two ongoing study was unclear and they are summarised in section 4.6.

Reasons for exclusion are described in Appendix 2. Of the studies that were potentially relevant to the symptomatic population, the main reasons for exclusion were no eligible AI software or AI not used in conjunction with radiology specialist (n = 119), or population not referred from primary care (n = 30). Only one identified study was conducted in a population referred from primary care, however the comparison was not relevant (AI software alone versus radiologist alone).[21] Of the studies potentially relevant to the incidental population, the main reasons for exclusion were no relevant outcome (n = 45), or no eligible AI software (n = 28).

As described in Section 3.1.7, to provide the closest available evidence to that required in Table 1, we looked for excluded studies that had (1) eligible AI software, and (2) compared radiology specialist in conjunction with AI software to radiology specialist alone, but where the referral status of the population was unclear. Studies that had an explicitly excluded

population (for example, a health screening population, pre-operative CXR, inpatients, A&E) remain excluded. Six such studies were identified (Table 2).

Figure 1: PRIMSA flow diagram



4.2. Study characteristics and key biases of selected excluded studies

Characteristics of the summarised studies are described in Table 2. In brief, six studies were summarised[22-26] (Siemens 2022, unpublished AIC submission from Siemens Healthineers), all of which employed a retrospective study design. Four studies were published; two were provided by the companies and not peer reviewed, one of these is preprint[22] and the other is ongoing (Siemens 2022). The studies were carried out in the USA, [26] (Siemens 2022) Germany,[26] Korea,[23-25] and the UK.[22]

CXR images were obtained from hospital databases,[22-25] the Lung Image Database Consortium,[26] a health centre database,[26] or from [redacted] (Siemens 2022). The number of CXR images included in the studies ranged from 100[26] to 434[25] and the number of participants who provided CXR data ranged from 100[26] to 400[22] (not reported in Siemens 2022). No information was provided about the referral route of patients who provided CXR data in any of the studies. It is plausible that the studies include both symptomatic patients and those who had CXR for reasons unrelated to lung cancer, as well as those from excluded populations such as people referred from other health care settings.

The characteristics of the CXRs assessed by the studies differed both within and across studies (Table 2). The UK study[22] identified random samples of patients who had a clinical text report indicating potentially malignant CXR and a follow-up CT, and those with a clinical text report of no urgent findings. Nam 2020 and Jang 2020 [23, 24] both included a large proportion of confirmed cancer cases with false-negative CXRs prior to diagnosis. Homayounieh[26] and Siemens 2022 selected CXRs to ensure negative and positive cases with different levels of difficulty in detection were included. Koo 2021[25] included adults with three or fewer nodules on both CXR and CT with at least one nodule pathologically confirmed on biopsy as either benign or malignant.

Images were assessed by a mix of consultant radiologists, board-certified radiologists, radiology trainees and reporting radiographers,[22] experienced radiologists, [23] experienced radiologists and radiology residents,[24, 25] and senior and junior radiologists.[26] The experience of the radiologists was not reported in one study (Siemens 2022). The readers had one[24] to 35[26] years of experience of reporting CXR.[22-26].One study reported the number of readers with fewer or more than 4 years of experience (Siemens 2022). The number of clinicians

included in the studies ranged from four[23, 25] to 11.[22] The accuracy of readers in detecting nodules or lung cancer with and without AI software was each compared with a ground-truth or reference standard, and these varied between the studies (Table 3). The threshold for defining a positive index test result (i.e. what was considered to be a nodule on CXR) was not defined in any of the studies.

Three studies assessed Lunit INSIGHT,[23-25] one assessed Red Dot Behold.ai,[22] and two assessed AI-Rad Companion Siemens.[26] (Siemens 2022) It is unclear whether the prototype AI software described in Siemens 2022 is commercially available.

Only a small number of outcomes that are relevant to the present review were assessed: test accuracy (lung cancer),[22] test accuracy (lung nodules), [23-26] (Siemens 2022) CT referrals,[22, 24] acceptability of AI to clinicians,[22] and CXR reading times.[24, 25]

The following risks of bias and applicability concerns were present in the reviews:

- Retrospective study designs were used in all of the studies. There is therefore the potential for selection bias, missing data and confounding.
- In all of the summarised studies, assessments were conducted on test-sets of data interpreted outside clinical practice. Caution is needed in extrapolating from these types of studies as prior evidence suggests little-to-no association between performance in this environment and that seen in clinical practice.[27]
- Only one study was conducted in the UK,[22] however it is unclear if the population the CXRs were taken from are reflective of people who would be referred from primary care in a real world setting. The generalisability of results from the other five studies is similarly limited in this way, and also because populations from USA and Korea may differ to the UK population in disease prevalence rates, age and comorbidities, and ethnic diversity.[28] There may also be differences in treatment settings and in the training and expertise of radiologists.
- AI software manufacturers were involved in three of the six studies (financial support n = 2[26] (Siemens 2022), employees authors n = 1[22]). Prior evidence suggests that studies conducted by drug/device manufacturers tend to report more favourable results than non-industry studies. [29] Caution in interpretation of these studies is warranted until independent assessment of the AI software is obtained.

- Each radiologist interpreted each CXR with and without AI software. In three studies [22, 24] (Siemans 2022) there was a washout period between readings, whereas in others [23, 26] the radiologist was aware of their initial decision at the second reading. This is not reflective of UK clinical practice and there is concern that the first reading could influence the second reading.
- The threshold for defining a positive index test result was not defined in the studies, therefore it is not possible to know whether the results of these studies are reflective of how AI would perform under clinical practice conditions, nor is it possible to know whether the results are comparable between studies.
- Where CT referrals were reported, these were hypothetical referrals rather than actual referrals and may not reflect real-world practice.

Table 2: Characteristics of summarised (but ineligible) studies

Study details	Population	Interventions	Notes
Dissez 2022[22] UK Retrospective cohort study, single centre Database dates: 2020 Aim: To evaluate the impact of an AI algorithm in augmenting the ability of clinicians to identify lung cancer on CXR No funding. Several authors are employed by and/or have stock/stock options in Behold.ai	400 CXRs taken from 400 adults with either: - clinical text report indicating potentially malignant CXR and follow-up CT (random sample of n=200) - clinical text report of no urgent findings (random sample of n=200) 132/400 CXR ground-truthed as suspicious for lung cancer 72/400 CXR clinically confirmed lung cancers CXRs taken from hospital databases	**Red Dot (Behold.ai) + radiologists** **Comparator:** 11 clinicians (3 FRCR consultant radiologists, 2 board-certified radiologists, 2 radiology trainees, 4 reporting radiographers) not involved in ground-truthing Two sessions with 4-week washout: CXRs reviewed without AI assistance in first session and with AI assistance in second image. Clinicians provided with basic clinical information, including age and sex **Reference standard:** lung cancer diagnosis (clinically confirmed outcomes collated by radiologist including repeat CXR and CT outcomes, lung cancer diagnosis, TNM staging and biopsy outcomes) **Comparison of relevance** extracted here is average accuracy of radiologists in identifying lung cancer (versus reference standard) and average performance of radiologists + AI software in identifying lung cancer (versus reference standard)	Referral route unclear, not known if symptomatic or incidental. Population from retrospective CXR collected in one UK NHS hospital during 2020. Participating clinicians had a range of 1 to 18 years' experience. Each of 11 clinicians reviewed each x-ray with and without AI, unclear if 4-week washout is sufficient. CT referrals were hypothetical rather than actual, as CXRs were retrospectively selected from databases. Full details and responses of the clinician survey not reported.
Nam 2020[23] Korea	218 CXRs from 218 people with pathologically confirmed lung cancers at percutaneous lung biopsy	**Lunit INSIGHT version 1.0.1.1 + radiologists** **Comparator:** four experienced thoracic radiologists not involved in ground truth	Referral route unclear, not known if symptomatic or incidental.

Retrospective cohort study, single centre Database dates: 2017 to 2018 Aim: To evaluate a deep learning–based algorithm for detecting lung cancers not reported on CXR Non-commercial funding	from single hospital AND false-negative posteroanterior CXR prior to biopsy (n=168) Nodules <5mm excluded Normal true-negative CXR confirmed on same day CT (n=50) CXRs taken from hospital database	Each reader reviewed CXR and made judgement (test 1), then reviewed results of algorithm and initial decision, and modified decision (test 2) **Ground truth**: CT (3-38 days from CXR) and/or re-evaluation of CXR reviewed by two experienced thoracic radiologists **Comparison of relevance** extracted here are average of radiologists (versus ground truth) and average of radiologist with AI (versus ground truth)	Korean population likely low generalisability to UK population. Population mainly confirmed lung cancer and false negative CXR prior to biopsy, some with true negative CXR. Nodules smaller than 5 mm were excluded Reader aware of initial decision on second read with algorithm Experience of radiologists ranged between 5 and 9 years Readers were aware of the characteristics of the CXRs but not the proportion of positive to negative cases
Jang 2020[24] Korea Retrospective cohort study, single centre Database dates: 2010 to 2014 Aim: To evaluate the efficacy of a deep learning–based automatic detection algorithm in	351 CXRs taken from 351 people diagnosed with lung cancer at a single tertiary hospital AND visible cancer on prior CXR at least 3 months before diagnosis when reviewed retrospectively by radiologist (n=117): - Detected without misinterpretation n=12 - Overlooked cancers n=105 (detected with misinterpretation n=23, undetected n=82)	**Lunit INSIGHT version 1.2.0.0 + radiologists** **Comparator**: six experienced thoracic radiologists and three radiology residents not involved in reference standard Each reader reviewed each CXR twice, once with and once without algorithm, with ≥4 week interval between sessions **Reference standard**: lung cancer lesion areas identified on CXRs and CT scans at time of diagnosis marked in consensus by two authors	Referral route unclear, not known if symptomatic or incidental, population is people with lung cancer and cancer visible on CXR prior to diagnosis, control group is those with normal CXR Korean population likely low generalisability to UK population. Experience of radiologists ranged between 1 and 12 years, radiology residents were either 2nd or 3rd years Observers were blind to clinical information

observer performance for detection of lung cancers on CXR Non-commercial funding	Healthy control with normal CXR confirmed at CT (n=234) CXRs taken from hospital database	**Comparison of relevance** extracted here are average accuracy of radiologists (versus reference standard) and average of radiologists with AI (versus reference standard	A web-based tool was used to document the readers results and calculate agreement between the AI and non-AI reading of the CXRs CT referrals were hypothetical rather than actual, as CXRs were retrospectively selected from databases
Koo et al 2021[25] Korea Retrospective cohort study, single centre Database dates: 2016 to 2018 Aim: To assess a deep convolutional neural network algorithm for pulmonary nodules on CXR Non-commercial funding	434 CXRs from 378 adults from a tertiary hospital with ≤ 3 nodules on both CXR and CT with ≥ 1 nodule pathologically confirmed on biopsy as either benign (n=246) or malignant (n=132) and nodules evident on chest CT visible on CXR CXRs taken from hospital database	**Lunit INSIGHT CXR version 1.00 + radiologist** **Comparator:** Two radiology residents and two thoracic radiologists Each reader reviewed CXR without AI and then ≥ 3 weeks later re-evaluated with the AI data **Reference standard:** Consensus from two thoracic radiologists with 10 and 7 years of experience using CR or CT **Comparison of relevance** extracted here are average accuracy of radiologists (versus reference standard) and average of radiologists with AI (versus reference standard)	Referral route unclear, participants from hospital setting in Korea. Likely low generalisability to UK population. Radiologist expertise differed (between 7 and 10 years of thoracic imaging experience for 2 radiologists, and 2 and 4 years of experience for 2 radiological residents), results for overall group extracted only Readers were blind to clinical information but were aware that CXRs would exhibit more nodules than CXRs from a normal clinical setting but not how many CXRs featured nodules Unclear if radiologists had their original decisions at the second reading
Homayounieh et al 2021[26] USA and Germany	100 CXRs taken from 100 adults with posterior-anterior CXRs taken between 2000 and 2010 (n=25 with absence of any abnormality, n=50	**AI-Rad Companion Chest X-ray (Siemens Healthineers) + Radiologist** **Comparator:** Radiologist alone (7 staff radiologists and 3 radiology residents)	Referral route unclear and generalisability to a UK primary care referred population unclear Radiologist expertise differed widely (between 2.5 years and 35 years for staff radiologists; radiology residents were in first year)

Retrospective cohort study, Two centres Database dates: 2000 to 2010 Aim: To assess the ability of an AI algorithm to detect pulmonary nodules from CXR Commercial funding: (Lunit Inc; Riverain Technologies Inc; Siemens Healthineers AG)	presence of pulmonary nodules of varying detection difficulties (20 challenging, 7 moderate, 23 easy), n=25 non-nodular abnormalities) CXRs taken from two databases, an ambulatory health care centre and the Lung Image Database Consortium	One month period between readings without AI first, and then original decisions available at the second reading **Ground truth:** consensus from 2 thoracic radiologists (with 14 and 16 years of experience respectively) **Comparison of relevance** extracted here are average of radiologists (versus ground truth) and average of radiologist with AI (versus ground truth)	Radiologists had their original decisions at the second reading (although described as a washout period) CXRs were selected to ensure negative and positive cases and the detection level varied (nodule sizes between 4 to 28 mm). Readers were aware that there were positive and negative CXRs but not the ratio of positive or negative CXRs The order CXRs were read was randomised across readers and reading sessions and findings recorded on an electronic case record form. One test reader was excluded from the analysis as did not follow the exact instructions
Siemens (Siemens Healthineers AIC submission) [redacted] Database / recruitment date: not reported Aim: to assess the use of AI for	[redacted] CXRs taken from [redacted]	**Prototype AI Rad Companion Chest X-ray algorithm (Siemens Healthineers) + Radiologist** **Comparator**: Seven radiologists Each radiologist assessed CXR unaided and with the AI after a four-week washout **Ground truth:** Two thoracic radiologists using CXR and CT **Comparison of relevance** extracted here are average accuracy of radiologists (versus reference standard) and average of radiologists with AI (versus reference standard)	Ongoing study with limited detail of early results provided Referral route unclear Radiologist expertise differed (four with > 4 years' experience and 3 with <4 years' experience) Unclear if prototype is commercially available AI Generalisability to a UK primary care referred population unclear

detecting pulmonary nodules and masses on CXR with accompanying chest CT Commercial funding (Siemens Healthineers)			

4.3. What are the test accuracy and test failure rates of adjunct AI software to detect lung cancer on CXR?

We did not identify any studies that met the inclusion criteria for this question.

Results of six summarised (but ineligible) studies are reported in Table 3. Studies reported test accuracy for individual readers and/or mean values for all readers; the data summarised in Table 3 are the mean values across readers.

One study examined the test accuracy of AI software to detect lung cancer on CXR.[22] In this UK study of Red Dot (Behold.ai), sensitivity was significantly higher for the interpretation of CXR with AI (77%, 95% CI 75% to 80%) than without AI (66%, 95% CI 59% to 71%). No difference was observed for specificity (see Table 3).

Five studies examined the test accuracy of AI software to detect lung nodules on CXR.[23-26] (Siemens 2022) Three studies from Korea[23-25] assessed different versions of the Lunit INSIGHT AI software. No statistically significant differences were observed in sensitivity or specificity between readers with or without AI in the studies by Nam 2020[23] and Jang 2020[24] (see Table 3). In the third paper [25], assessment of test accuracy was conducted for *any* nodule and *each* nodule. In the analysis of any nodule, sensitivity was 95.1% for readers with AI software and 92.4% without AI software, and specificity was 97.2% for readers with AI software and 93.1% without AI software. In the analysis of each nodule, sensitivity was 93.9% for readers with AI software and 88.6% without AI software. Specificity was not reported. Instead, false positive rates were reported to be 3.2% for readers with AI software and 6.3% without AI software. Caution is required in the interpretation of false positive data as the paper reports that the false positive rate is the total number of false positives divided by the number of CXRs which is a non-standard calculation. It is not possible to know if the above estimates reflect true differences between assessment with/without AI software as no statistical analyses were presented in the paper for any of the above test accuracy metrics, and there was insufficient data to allow us to conduct our own analyses.

Two studies (Hamayounieh 2021[26] and Siemens 2022) assessed versions of the Siemens Healthineers AI-Rad Companion Chest X-ray in USA and German populations (one study) or in [redacted] populations alone (Table 3). No statistically significant differences were observed in

sensitivity or specificity between radiologists with or without AI software in the studies by Homayounieh 2021[26] and Siemens 2022.

One study[22] reported mean number of cancers detected and found no significant differences with and without AI software (54 cancers, 95% CI 42 to 59; and 46 cancers, 95% CI 38 to 51, respectively).

None of the six studies reported AI software test failure.

Table 3: Test accuracy results from summarised (but ineligible) studies

Study name	AI name	No. of Patients	No. of CXR	No. of cancers / nodules	Group	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lung cancer detection											
Dissez 2022 [22]	Red Dot (Behold.ai)	400	400	72	With AI	NR	NR	NR	NR	77% (75% to 80%)	75% (71% to 77%)
					Without AI	NR	NR	NR	NR	66% (59% to 71%)	81% (77% to 85%)
Nodule detection											
Nam 2020[a] [23]	Lunit INSIGHT version 1.0.1.1	NR	NR	NR	With AI	357	36	315	164	53% (49% to 57%)	82% (77% to 87%)
					Without AI	316	44	356	156	47% (43% to 51%)	78% (72% to 84%)
Jang 2020[a] [24]	Lunit INSIGHT version 1.2.0.0	351	351	117	With AI	66	19	51	215	56% 47% to 65%)	92% (88% to 95%)
					Without AI	50	24	67	210	43% (34% to 52%)	90% (86% to 94%)
Koo 2021 [25] - per patient any nodule	Lunit INSIGHT version 1.0.0.0	378	434	165	With AI	NR	NR	NR	NR	95% (NR)	97% (NR)
					Without AI	NR	NR	NR	NR	92% (NR)	93% (NR)
Koo 2021 [25] - per nodule		NR	NR	NR	With AI	NR	NR	NR	NR	94% (NR)	NR (NR)
					Without AI	NR	NR	NR	NR	89% (NR)	NR (NR)
Homayounieh 2021 [26]	AI-Rad Companion Chest X-ray	100	100	NR	With AI	26.4	2.5	23.6	47.5	55% (48% to 63%)	95% (91% to 9%)
					Without AI	23.6	4.1	26.4	45.5	45% (38% to 53%	93% (89% to 96%)
Siemens 2022	Prototype AI Rad Companion Chest X-ray algorithm	1018	1018	NR	With AI	NR	NR	NR	NR	[redacted](NR)	[redacted](NR)
					Without AI	NR	NR	NR	NR	[redacted](NR)	[redacted](NR)

[a] 95% confidence intervals calculated by the EAG using TP, FP, FN, and TN. Data are mean values for all readers.

4.4. What are the practical implications of adjunct AI software to detect lung cancer on CXR?

We did not identify any studies that met the inclusion criteria for this question.

Two of the summarised (but ineligible) studies provided information on the potential referrals for CT.[22, 24] No statistically significant differences were observed in the number of people who might be recommended for CT follow-up between readers with and without AI: Red Dot (Behold.ai) 144/400 (36%) (95% CI 119 to 172) potential referrals with AI and 117/400 (29%) (95% CI 93 to 147) potential referrals without AI;[22] Lunit INSIGHT: 96/351 (27%; 95% CI 22.8 to 32.3 calculated by EAG) with AI and 80/351 (23%; 95% CI 18.5 to 27.5 calculated by EAG) patients without AI.[24] It is important to note that these are hypothetical referrals. We found no evidence on the impact of AI on the readers behaviour in real-world clinical practice.

Two studies[24, 25] reported information on reading times. No statistically significant differences were observed in average image reading times between readers with and without AI: Siemens Healthineers AI-Rad Companion 22.5 (SD 40.3) seconds with AI, 24.3 (SD 27.4) seconds without AI, per image;[24] Lunit Insight 171 (SD 33.8) minutes with AI, 211.25 (SD 38.4) minutes without AI, to read 434 CXR.[25]

One study reported on the acceptability of Red Dot (Behold.ai) amongst ten out of 11 study clinicians.[22] Eight clinicians indicated that reporting was not slowed down by AI, and nine stated that "the heatmaps [visual display of findings suspicious of lung cancer on CXR] produced by the AI model were helpful to understand the algorithm's attention points." (Dissez 2022, p8)[22]

4.5. What is the clinical effectiveness of adjunct AI software applied to CXR?

We did not identify any studies that met the inclusion criteria for this question. None of the six summarised (but ineligible) studies reported clinical effectiveness outcomes.

4.6. Summary of ongoing trials

No ongoing trials meeting the inclusion criteria were identified. As described in section 3.1.7, we looked for ongoing trials assessing eligible comparisons.

We identified one ongoing trial (KCT0005466) comparing Lunit INSIGHT in conjunction with a radiologist versus radiologist alone, however the population is those undergoing CXR for any reason in the outpatient department. It is not known whether the participants had CXR for symptoms due to cancer or for reasons other than cancer, or if they were referred from primary care.

Details of one ongoing study (NCT05489471), identified from the Lunit company submission, are unclear. The proportion of GP referrals, accident and emergency attendances and in-patients is not known, the AI software is not named (but the study is funded by Lunit) and it is not clear whether the comparison is AI software in conjunction with a radiologist versus radiologist alone. This UK based study is currently not yet recruiting and has an estimated primary end date of July 2023.

In addition, the Siemens 2022 study provided in the Siemens Healthineers' company submission summarised above is currently ongoing.

Table 4: Ongoing studies

Trial identifier number	**KCT0005466**
Title of project	Prospective evaluation of deep learning-based detection model for chest radiographs in outpatient respiratory clinic
Trial completion date	31/05/2021 (no results posted)
Trial identifier number	**NCT05489471**
Title of project	A Study to Assess the Impact of an Artificial Intelligence (AI) System on Chest X-ray Reporting
Trial completion date	Estimated primary end date of July 2023.

5. Cost-effectiveness

5.1. Results of literature searches

1120 studies were identified through database searches (817 in Medline and 303 in Tufts CEA). Of these, 29 studies were retrieved for full text assessment (25 from Medline and 4 from Tufts CEA). These covered a wide range of methodologies and research questions. Reference lists of these studies returned four further studies of relevance to this review.

We did not identify any cost-effectiveness studies that directly compared CXR review by radiology specialist with adjunct AI and radiology specialist review without. However, two economic evaluation studies from the database search[30, 31] and an updated analysis of one of these[32] found through an authorship search were identified as useful to inform modelling techniques and parameter input sources. Similarly, four studies (one from the database search[33] and three from author searches[34-36]) provided detailed information on radiological and clinical pathways to lung cancer diagnosis in the UK. A systematic review and meta-analysis on the diagnostic performance of CXRs in symptomatic primary-care populations[37] was also retrieved from the search.

These studies were retained and summarised narratively to include information of relevance to populate the conceptual model. No formal data extraction or quality appraisal was conducted. The studies by Snowsill et al. (2018)[31] and the Exeter Test Group and Health Economics Group (2022)[32] were not summarised. Information for the diagnostic component of the conceptual model was prioritised due to project time constraints, whereas these studies[31, 32] pertained more to the longer-term treatment costs and utilities.

5.2. Description of the evidence.

Bajre et al. (2017)[30]

Bajre and colleagues used a decision tree structured model to assess the cost-effective of trained reporting radiographers compared with radiologists for the reporting of CXR in people suspected of having lung cancer.[30] The model simulated a pathway for a hypothetical cohort of 1000 people undergoing CXR for suspected lung cancer, with cost-effectiveness calculations concluding at five years. The model started with a cohort of people receiving

either a radiologist-reported CXR or radiographer-reported CXR. The pathway for both strategies were the same. Proportion of those with true disease status was known, characterised by the prevalence of lung cancer. People with lung cancer who had a positive CXR result received a confirmatory test of a CT scan, which also provided staging. The authors included stages I, II, II and IV lung cancers. People with a false negative result presented later to the Accident and Emergency department, where there were diagnosed with lung cancer and staged. People who had a false positive result following CXR received a CT scan that confirmed no lung cancer was present. People with no lung cancer and who had been correctly identified as negative by the CXR received no further testing/imaging.

Information required to populate the model was obtained from the literature and NHS reference costs. The model required information about the prevalence of lung cancer, sensitivity and specificity of radiologist-reported and radiographer-reported CXR to identify lung cancer, as well as sensitivity and specificity for radiologist-reported CT scan to confirm lung cancer diagnosis and probabilities. Though not explicitly stated, confirmatory diagnosis was made by the radiologist. The proportion of people diagnosed at first presentation were obtained from statistics published by Cancer Research UK in 2013 (reference number 32 in Bajre et al. (2017)[30]). Additionally, information was required about the probability of lung cancer by stage at second presentation following misdiagnosis. All costs included in the model were reported in 2014/15 prices. Costs were required for radiologist and reporting radiographer reading of CXR, cost of CT scans, and total costs of treatment by stage. Authors were not explicit about which treatment people received. The benefit of the strategies was reported in terms of cases detected at first presentation and quality-adjusted life years (QALYs) yielded. Utility values by stage of diagnosis were obtained from Naik et al., 2015.[38]

Several simplifying assumptions were made to have a workable model structure (Bajre et al., 2017 pg. 275):

- *Time taken to report CXR is 2 minutes for both reporting radiographers and radiologists*
- *False negatives present at A&E at a later date, at which point disease may have advanced a stage (for patients at stage I to III)*

- *Sensitivity and specificity of radiographer reporting of CXR and radiologist reporting of both CXR and CT-scan is independent of disease stage or other patient characteristics such as age*
- *QOL in the year following diagnosis (according to stage at diagnosis) is maintained in subsequent years*
- *There is no QOL impact arising from false positive reporting*
- *Findings for non-small cell lung cancer are representative for lung cancers in general*

The perspective and setting of the economic analysis were not clearly defined but it appears to be from the NHS and Personal Social Services (PSS) in a secondary care setting, based on the cost inputs. The results of the analysis were presented in terms of an incremental cost-effectiveness ratio (ICER), expressed as cost per QALY. The authors undertook probabilistic sensitivity analysis (PSA) to assess the joint uncertainty in key model input parameters: prevalence of lung cancer, sensitivity and specificity of radiologist and radiographer reporting of CXR, lung cancer stage distribution at initial CXR and stage progression following misdiagnosis. Authors stated the sampling distributions for the parameters included in the PSA but have not reported their parameters. The authors undertook threshold analysis but not one-way sensitivity analysis.

Authors reported disaggregated results for both strategies. Results were reported on the number of people expected to be diagnosed with lung cancer, QALYs yielded, and treatment costs, all by stage. The QALYs yielded appeared to be high, with stage IV expected to yield more QALYs that stage III and II, respectively. There were modest QALY gains by strategy and by stage, with stage I having the greatest expected gain of 2.4 QALYs, favouring radiographer reporting. Radiographer reporting yielded more overall QALYs, but it was unclear with the inputs reported why stage II and IV the radiologist reporting QALYs was greater for stages II and stage IV. Radiographer reporting diagnostic and treatment costs were cheaper than radiologist reporting costs. Overall results showed that radiographer reporting of CXR dominated radiologist reporting. PSA results showed that radiographer reporting continued to dominate radiologist reporting in 98% of the iterations. Based on the model structure, it's inputs and assumptions, the authors concluded that the use of trained reporting radiographers to report CXR is cost-effective and an increased role for reporting

radiographers in the would be beneficial to meet hospital waiting time targets for lung cancer diagnosis.

Foley et al. (2021)[33]

Foley and colleagues conducted a retrospective review of Trust audit data (Royal United Hospitals Bath NHS Foundation Trust) to analyse the use of CXR as the first-line investigation in primary care patients with suspected lung cancer. 1,488 of the 16,495 primary care referrals received between 1st June 2018 and 31st May 2019 were for suspected lung cancer. CXRs were coded by result as CX1, normal but a CT scan is recommended to exclude malignancy; CX2, alternative diagnosis; or CX3, suspicious for cancer. Outcomes for the study cohort were stratified by CX code and included patient characteristics, number undergoing CT scan, number of lung cancers diagnosed, stage at diagnosis, time from initial CXR to CT scan, time from CT request to CT scan, time to diagnosis, treatment strategy taken and mortality (over an average follow-up period of 322 days in the total cohort). See Table 5 for results of key outcomes.

Table 5: Outcome data reported by code (Foley et al. 2021)[33]

Outcome	CXR report code			Statistical significance (p < 0.05)
	CX1 (normal but CT scan recommended to exclude malignancy)	CX2 (alternative diagnosis)	CX3 (suspicious for malignancy)	
Total number of CXRs (%)	1,056 (75)	288 (20)	72 (5)	-
Number referred for CT (%)	107 (10)	107 (37)	66 (92)	-
Number of lung cancers diagnosed (%)	10 (1)	29 (10)	49 (68)	-
Number diagnosed at advanced stage IIIc/IV (%)	5 (50)	11 (38)	28 (57)	(p = 0.26)
Number of days from CXR to CT*	34.6	19.6	1.9	(p < 0.001)
Number of days from CXR to diagnosis*	89.7	65.3	30.2	(p < 0.001)
Number receiving treatment with curative intent (%)	4 (40)	14 (48)	13 (27)	(p = 0.14)
Number of deaths in follow up period (all-cause mortality) (%)	5 (50)	10 (34.5)	27 (55.1)	(p = 0.42)
CXR, chest x-ray; CT, computed tomography; *mean value				

Based on these findings, authors concluded there was significant delay in lung cancer diagnosis in patients who received a CX1 'normal' initial CXR result (p < 0.001) and the majority of patients with a 'normal' or 'abnormal' CXR are diagnosed at an advanced disease stage (p = 0.26) with no difference in survival outcomes based on the CXR findings (p = 0.42).[33]

Bradley et al. (2021)[34]

Bradley and colleagues undertook a retrospective observational study using routinely collected healthcare data from Leeds Teaching Hospitals NHS Trust (LTHT).

All patients diagnosed with primary lung cancer between January 2008 and December 2015 with a GP-requested CXR in the year before diagnosis were coded based on the result of the earliest CXR in that period. CXR report codes were assigned: 1. Suspicion of lung cancer identified/urgent investigation needed; 2. Abnormality identified/non-urgent investigation indicated, including diagnoses of pneumonia or consolidation even if repeat imaging was not explicitly suggested; 3. Abnormality identified but no further investigation/assessment indicated; and 4. Normal CXR, no abnormalities identified.

The sensitivity of CXR was calculated and analyses performed on time to diagnosis, stage at diagnosis and survival outcomes. Statistical analysis on these outcomes was performed by combining CXR codes 1 and 2 together to form a 'positive' result group and codes 3 and 4 to form a 'negative' result group. However, authors present numerical outcome data for all codes separately as well as combined groups. See Table 6 for summary of key data by individual codes.

Table 6: Key outcome data reported by individual codes (Bradley et al. 2021)[34]

Outcome	**Initial CXR code**				**Total**
	1	**2**	**3**	**4**	
Number of CXRs (%)	1383 (65)	370 (17.4)	230 (10.8)	146 (6.9)	2129
Time from CXR to diagnosis, median days (IQR)	36 (23, 63)	93 (55, 154)	211 (181, 296)	193 (87, 279)	51 (29, 107)
Survival from CXR, median days (IQR)	313 (126, 877)	400 (163, 964)	408 (238, 958)	420 (214, 1117)	345 (148, 920)
Stage I/II at diagnosis, n(%)[95% CI]	397 (28.7) [26.4 to 31.2]	111 (30) [25.4 to 35.0]	83 (36.1) [30.0 to 42.7]	43 (29.5) [22.4 to 37.7]	634 (29.8) [27.9 to 31.8]
Stage III/IV at diagnosis, n(%)[95% CI]	981 (70.9) [68.4 to 73.3]	259 (70) [65.0 to 74.5]	147 (63.9) [57.3 to 70.1]	103 (70.5) [62.4 to 77.7]	1490 (70) [68.0 to 71.0]
Stage unknown at	5 (0.4)	0	0	0	5 (0.2)

diagnosis, n(%)[95% CI]					

Data were also presented on the number of people who had further CXRs requested by their GPs, with median time to second CXR and median times to diagnosis from initial CXR. Of 376 patients with an initial CXR that was 'negative' (codes 3 and 4), 98 (26.1%) had at least one further CXR. Sensitivity calculated based on initial CXR (codes 1 and 2) was 82.3% (95% CI = 80.6% to 84.1%).

Authors concluded that sensitivity results supported previous systematic review findings,[39] and whilst those with a 'positive' initial CXR finding had a median of 43 days to diagnosis compared with 204 days for those with 'negative' findings, no direct association with time to diagnosis was found between stage at diagnosis or survival in this study.

Woznitza et al. (2018)[35]

Woznitza and colleagues conducted a four-month feasibility study (November 2016 to March 2017) at a single radiology department at an acute general hospital (Homerton University Hospital, London). The primary outcome was to establish the feasibility of an immediate reporting service for CXRs. Comparison between CXR referrals from general practice that received an immediate and routine report was made to determine the number of lung cancers diagnosed, time to diagnosis, time to CT and number of urgent referrals to respiratory medicine.

From 1,687 CXRs of people referred from general practice over the 26-week study period, 36 patients (22 immediate CXR report, 14 routine CXR report) had a CT scan arranged by radiology following a suspicious CXR. This equated to less than one additional unplanned patient per week (mean 0.8 scans per week) accommodated by the CT department. Time from CXR to CT was shorter in the immediate report group with a mean of 0.9 days (SD=2.3) compared to routine reporting 10.6 days (SD=4.5) ($p>0.0001$). No apparent difference was found for time to discussion at MDT.

The study also gave detailed description of the radiology department demographics and processes for reporting and referral. Results of all CXRs included in the study and pathways taken were explained, including 17 patients with a normal or non-cancer diagnosis at CXR who were subsequently diagnosed with lung cancer.

The authors concluded that it was feasible to introduce a radiographer-led immediate CXR reporting service but a definitive study assessing outcomes would be needed to determine whether this would have an impact on mortality and morbidity for patients.

Woznitza et al. (2022)[36]

Woznitza and colleagues conducted a prospective, block-randomised controlled trial (RadioX) at a single acute district general hospital in London (Homerton University Hospital). People referred for CXR from primary care attended sessions that were pre-randomised to either immediate radiographer (IR) reporting, or standard radiographer (SR) reporting within 24-hours. Those who received SR reporting were the control group as this was usual practice in the department. In the intervention group, CXRs were reported whilst the patient was still in the department with all patients with CXR findings suspicious for lung cancer offered a same day CT scan. Those who declined were scheduled for another day.
8682 CXRs were performed between 21st June 2017 and 4th August 2018, 4096 (47.2%) for IR and 4586 (52.8%) for SR. Lung cancer was diagnosed in 49 patients. See Table 7 for summary outcome data from trial reporting arms.

Table 7: Outcome data reported by immediate and standard CXR reporting arms Woznitza et al. (2022)[36]

Outcome	**Immediate Reporting**	**Standard Reporting**
Total patients	4096	4586
Previous CXR, n (%)		
Yes	2297 (56.1)	2583 (56.3)
No	1799 (43.9)	2003 (43.7)
Previous CT, n (%)		
Yes	307 (7.5)	334 (7.3)
No	3789 (92.5)	4252 (92.7)
Lung cancer suspected, n (%)		
Yes	1326 (32.4)	1511 (33.0)
No	2757 (67.3)	3062 (66.7)
Known	13 (0.3)	13 (0.3)
Total cancers diagnosed, n (%)	27 (0.7)	22 (0.5)

2WW referral		
Yes	150 (3.7)	189 (4.1)
No	3946 (96.3)	4397 (95.9)
Time from CXR to diagnosis (days)		
Median (IQR)	32 (19, 70)	63 (29, 78)*
Mean (SD)	47.2 (35.8)	81.6 (78.5)
Time from CXR to discharge (days) (no cancer diagnosis)		
Median (IQR)	30 (17, 64)	27 (14, 61)
Mean (SD)	54.4 (60.4)	50.3 (63.7)
2WW, two week wait referral on suspected lung cancer pathway; *p=0.03		

Authors stated a health economic evaluation based upon their RadioX trial was to be reported separately.[36] The corresponding author was contacted and confirmed analysis of the data was still underway and they were unable to share any usable information at present (Nicholas Woznitza, Consultant Radiographer, University College Hospital London NHS Foundation Trust, personal communication 29.12.2022).

Dwyer-Hemmings & Fairhead (2021)[37]

Authors performed a systematic review of evidence to inform diagnostic accuracy of CXR to detect lung malignancy in symptomatic patients presenting to primary care. Nine databases were searched and data from included studies extracted to calculate sensitivity and specificity of CXR where possible. Risk of bias was assessed using a validated tool and random effects meta-analysis was performed. Ten studies were included. Sensitivity meta-analysis was performed in five studies which were not at high risk of bias, with summary sensitivity of 81% (95% CI: 74, 87%). Specificity could be calculated in five studies, with summary specificity of 68% (95% CI: 49, 87%). The authors concluded that the evidence for sensitivity was strong due to selection of studies that were not at high risk of bias, had low heterogeneity between studies, and low risk of publication bias. However, evidence for specificity was weaker due to heterogeneous study design and variance between reported outcomes.[37]

5.3. Clinical pathway for representation in model

The clinical pathway illustrated in Figure 2 was agreed in the NICE final scope.[10]

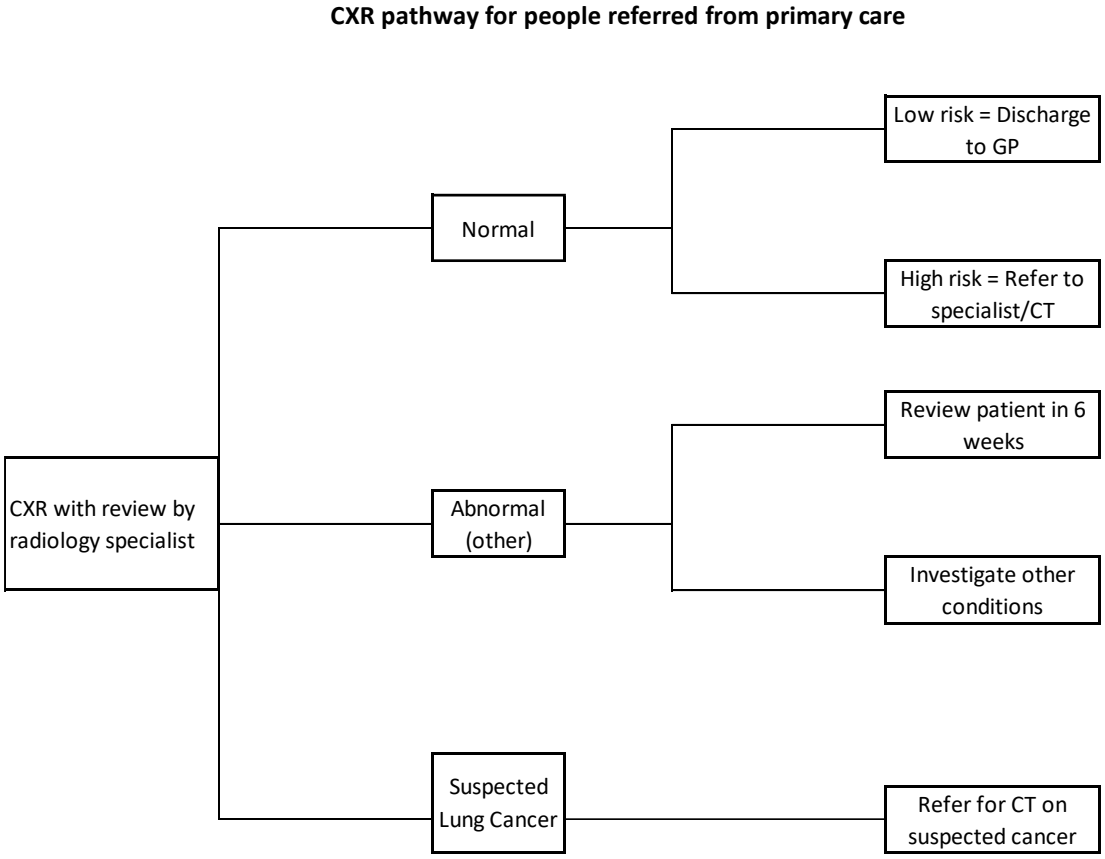


Figure 2: CXR pathway from NICE Final Scope[10]

Development of this pathway was supported by existing guidelines on the diagnostic and care pathway,[8, 9, 16, 17] and collaboration with specialist committee members (SCM) during the scoping process. Subsequent feedback from SCM and clinical experts generally supported this as a representation of the multiple pathways patients may follow after primary care referral for a CXR. All emphasised this was an aspirational pathway, with many alternative routes both in and out through to diagnosis, and was not particularly accurate in several Trusts.

When critical pathway events were mapped based on the early stages of the National Optimal Lung Cancer Pathway[17] using large cancer databases from two Trusts, 83 individual combinations of early pathway events in 1018 suspected lung cancer patients were found.[40] This highlights the complexity in defining a realistic structure on which to base the clinical component of an economic model. All models by their nature are a more simplistic format

of real practice. The balance is to represent the clinical pathway in sufficient detail to capture the main elements, whilst producing a model that is feasible to construct.

The availability of evidence to inform model parameters also influences the model structure. Where evidence is severely limited, a more simplistic model reduces the number of assumptions relied upon to achieve an executable model and reduces the uncertainty introduced.

Two studies identified in the literature search[33, 34] reported data for parameters which had the potential to support multiple differential pathways after CXR results, rather than just a lung cancer suspected and no lung cancer suspected route through model. However, there were limitations in how the data reported from both sources might be applied.

Overall, the EAG determined that the clinical pathway developed during the NICE scoping process was a realistic representation on which to base the conceptual model. Although concerns remained around feasibility of parameterising the model due to lack of available evidence and differences in outcome reporting, five differential pathways (A,B,C,D and E) were formulated with feedback from clinical experts and reference to the clinical guidelines.[8, 9, 16, 17]

Figure 3 shows where each pathway is situated, and each pathway is described in detail below.

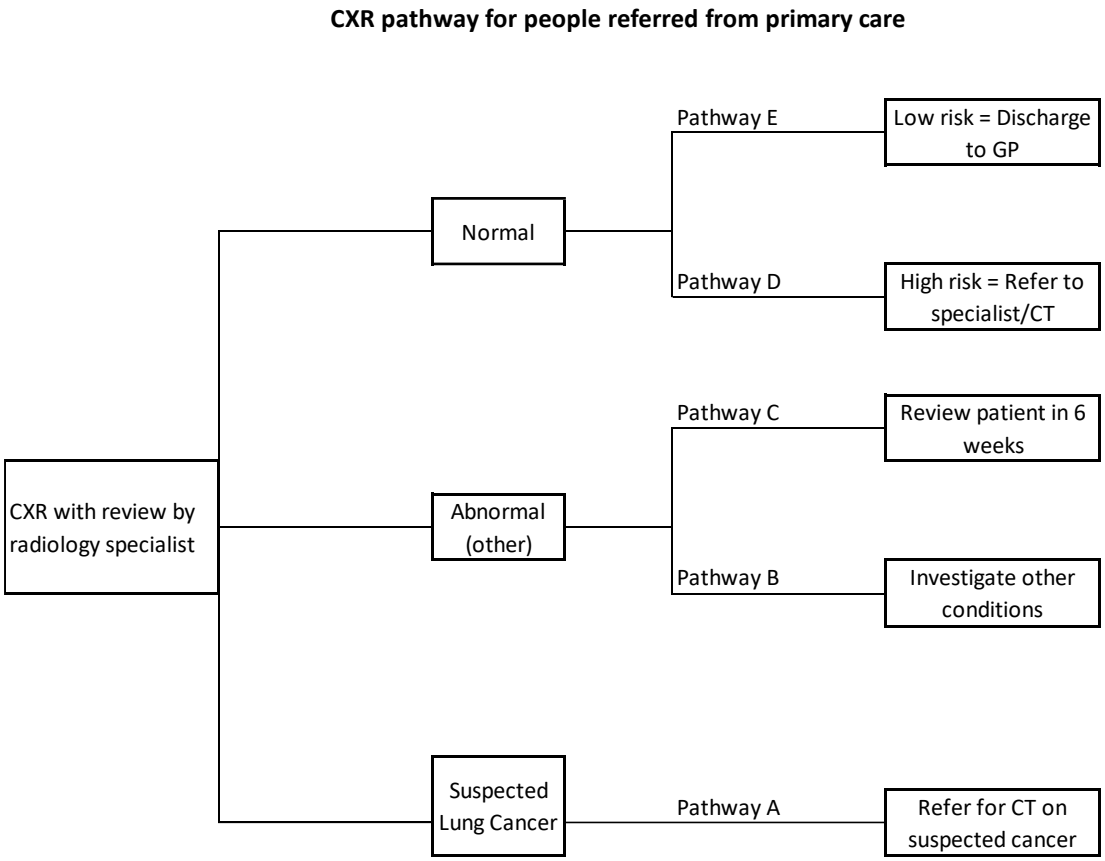


Figure 3: Clinical pathways for conceptual model

Pathway A

When CXR findings are suspicious for malignancy, referral for urgent CT on the suspected lung cancer pathway is made. There is a variation in practice across Trusts, but in many institutions highly suspicious CXR findings are flagged to secondary care lung cancer teams who request the CT scan and await referral to the suspected lung cancer clinic from the GP. Once reported CT scans are triaged by lung cancer team consultants. If they suggest probable lung cancer, an urgent lung cancer team appointment is arranged with appropriate tests e.g. spirometry, planned biopsy (EBUS) (Alberto Alonso, Consultant Radiologist, Manchester Hospital NHS Foundation Trust, personal communication 19.12.2022) for histopathological staging and to inform treatment options at the fast track lung cancer clinic.[16]

If the CT scan appears reasonably normal (despite CXR appearances) the lung cancer team write to the patient to inform them of their relatively normal CT appearances and arrange a non-urgent general respiratory clinic (not lung cancer clinic) outpatient appointment (Vidan Masani, Consultant Respiratory Physician and lead for lung cancer, Royal United Hospitals Bath NHS Foundation Trust, personal communication 1.02.2023). This also includes those who require investigation and management of pulmonary nodules according to BTS guidelines.[16, 17]

Pathways B and C

If CXR results are reported as 'abnormal' where findings are indeterminate or suggestive of an alternative diagnosis, people may follow pathways B or C. Here, findings are not sufficient to warrant further urgent investigation, but additional clinical enquiry is required.

Pathway B is taken when an alternative diagnosis is suspected and referral is made by the GP to a secondary care outpatient clinic with relevant expertise for that clinical finding e.g. non-urgent respiratory clinic.

Pathway C is followed when a six-week repeat CXR is advised in the report. The referral for repeat CXR is made by the GP and a radiologist or reporting radiographer compares the new image to the previous. If the abnormality is resolved, then no further action or follow up is required. If abnormal and suspicious, these cases are "red-alerted" or "upgraded" and the lung cancer team and referring GP are notified as per pathway A. The six-week repeat CXR is used in cases where there is need to exclude infection, try a course of treatment and reassess before considering CT (Jonanthan Rodrigues, Consultant Radiologist, Royal United Hospitals Bath NHS Foundation Trust, personal communication 13.12.2022).

Pathways D and E

Where CXR are reported as 'normal' findings may be unremarkable, but several Trusts (including Royal United Bath NHS Foundation Trust and Manchester University NHS Foundation) include an automatic caveat within the report 'please note that a normal CXR does not exclude malignancy. If there is still a strong suspicion of malignancy (weight loss/unresolved cough/significant or unresolved haemoptysis) referral for a CT scan is

advised'. This is to counter false reassurance in the case where clinical suspicion remains high.

People with normal results may therefore proceed along pathway D, where their GP considers them at high risk of lung cancer despite nothing detected on CXR and refers for CT scan and specialist review.

Pathway E is taken when the GP has no further concerns, no further diagnostic testing is requested, and management is continued under primary care.

5.4. Discussion of inputs to inform model structure

To formulate a final conceptual model an iterative process was used. This included identifying relevant intermediate and long-term outcome measures for parameterisation and selecting a structure which is most appropriate to support their inclusion.

This section describes available evidence, gaps in evidence and recommendations for appropriate evidence generation for a range of outcome measures. In this report these will be classified into intermediate measures (short to medium-term clinical outcomes encountered during the diagnostic process), long-term clinical outcomes, and cost inputs.

Intermediate measures for consideration:

- Accuracy to detect lung cancer

No eligible studies were found in the clinical-effectiveness review, but one of the six ineligible studies summarised examined the test accuracy of AI software to detect lung cancer on CXR.[22] In this UK study of Red Dot (Behold.ai), sensitivity was significantly higher for the interpretation of CXR with AI (77%, 95% CI 75%, 80%) than without AI (66%, 95% CI 59%, 71%). No difference was observed for specificity 75% (71%, 77%) with AI, 81% (77%, 85%) without (see Table 3).

A systematic review and meta-analysis identified in the cost-effectiveness literature review[37] provided evidence on the test accuracy of CXR to detect lung cancer in symptomatic patients presenting to primary care. In this population, specifically relevant to this review, summary sensitivity of 81% (95% CI: 74%, 87%) was calculated from five studies not at high risk of bias. Summary specificity of 68% (95% CI: 49%, 87%) was also obtained

from five studies but evidence was weaker due to their heterogeneous study design and variance between reported outcomes.[37] Findings of this systematic review were supported by two other studies from the cost-effectiveness search**.**[34, 39] A retrospective database study of all primary care referrals for CXR conducted by Bradley et al. 2021 reported sensitivity of 82.3% (95% CI = 80.6%, 84.1%). This was calculated based on an initial CXR coding system which included results suspicious for lung cancer and those with an abnormality identified but no urgent investigation indicated as a 'positive' result for CXR.

- Turnaround time (time from start of image review to radiology report)

Turnaround time (TAT) was identified in the final scope[10] as a potentially useful outcome measure in this assessment. From a modelling perspective, the placement in the pathway at which review time occurs is prior to the diagnostic decision outcome. This would be captured in a model as a resource use parameter used to calculate cost per image, where the rate of radiology specialist's pay is multiplied by length of time to review scan.

A reduction in cost may be expected where TAT is decreased. However, the direction and magnitude of this relationship is highly uncertain given the lack of evidence found on TAT with AI software assistance and the variation of estimates given for TAT without AI from the literature and clinical expert feedback.

Estimated TAT for CXR varies considerably. As discussed in section 4.4, of the ineligible studies reported on from the clinical search, two studies[24, 25] presented information on reading times. No statistically significant differences were observed in average image reading times between readers with and without AI: Siemens Healthineers AI-Rad Companion 22.5 (SD 40.3) seconds with AI, 24.3 (SD 27.4) seconds without AI, per image;[24] Lunit Insight 171 (SD 33.8) minutes with AI, 211.25 (SD 38.4) minutes without AI, to read 434 CXR.[25] which equates to an average of 23.6 secs per image with AI, and 29.2 seconds without (calculated by EAG).

No information was given on the methods used for timing. With regards to context, timings were recorded during specified reading sessions under study conditions, so how this would translate to reading times in clinical practice is unknown.

Methods by Royal College of Radiologists (RCR) to derive guidance on reporting output

figures are described comprehensively.[41] 80 reports for plain CXR per session (3 minutes per image) is the figure expected on average, over a six-month minimum period, per in-hours, on-site, non-acute four-hour reporting session in the NHS.[41]

SCMs advised average reading times of less than one minute up to five minutes, with an assumption of two minutes used in the economic evaluation by Bajre et al. (2017).[30]

Many factors impact on reporting output and are well-outlined by the Royal College of Radiologists (RCR), (2022).[41] Therefore, focus on this as an outcome measure, without appreciation of real-world context, is of little use unless a reduction in TAT can be shown to impact efficiency of workflow over a sustained period in the NHS environment. This needs to be considered when designing future studies.

Another anticipated benefit of reducing TAT is to increase the output of radiology specialists performing CXR reviews, thereby addressing the high demand for image reading and inherent limitations on workforce capacity. This is a potential value driver of AI software but would not be captured within the conceptual cost-effectiveness model. The potential value here would be recognised at a system level rather than at the patient level which is represented in the conceptual model.

- Technical failure rate

Technical failure rate was identified in the final scope[10] as a potential measure of interest. None of the six studies summarised in the clinical effectiveness review reported any information on technical failure rate in CXR.

- Impact of software output on clinical decision-making

Impact of software on clinical decision making is the primary measure of importance as the final CXR is result is determined by a radiology specialist, whether AI software is used or not. Even if the diagnostic accuracy of AI software alone is higher, the outcomes are mediated by human input. Results then determine which clinical pathway a patient will proceed down, affecting quantity and type of further tests.

No evidence was found on this, and the only extrapolated data was in the form of two studies[22, 24] that provided information on hypothetical referrals to CT. No statistically significant differences were observed in the number of people who might be recommended

for CT follow-up between readers with and without AI: Red Dot (Behold.ai) 144/400 (36%) (95% CI 119 to 172) potential referrals with AI and 117/400 (29%) (95% CI 93 to 147) potential referrals without AI;[22] Lunit INSIGHT: 96/351 (27%; 95% CI 22.8 to 32.3 calculated by EAG) with AI and 80/351 (23%; 95% CI 18.5 to 27.5 calculated by EAG) patients without AI.[24] It is important to note that these are hypothetical referrals as CXRs were retrospectively selected from databases in these studies. We found no evidence on the impact of AI on the readers behaviour in real-world clinical practice.

- Number of people referred for a CT scan

The number of people referred for a CT scan is dependent upon test accuracy and referral decision based on CXR result. As highlighted in the clinical review, evidence to inform these parameters which fall earlier in the clinical pathway was not available in the primary care population for CXR review with adjunct AI.

CT scans may be requested as a result of initial investigations, usually CXRs, undertaken in any of the pathways A, B, C, and D (see section 5.3 for detailed description). Therefore, proportion of people referred from each pathway for CXR would be needed for a model.

Only two studies identified in our literature search[33, 35] mentioned the number of people referred for a CT scan. Woznitza et al.[35] reported that a total of 36 patients out of the 1,687 referred for CXR from primary care underwent a CT scan. This included both suspected lung cancer and no-suspected lung cancer populations. Foley at al.[33] provided much more detailed information (see **Table 5**) and was specific to the GP-referred population with suspected lung cancer. The number and percentage of CT scans requested by the three CXR result codes were reported: CX3 (suspicious for malignancy), 92% (66/72) had a CT scan; CX2 (abnormal, alternative diagnosis), 37% (107/288); and CX1 (normal), 10% (107/1056) had a CT scan.

Whilst limited to only three potential pathways the data and reporting format from this paper[33] is useful to inform conceptual model parameters for current practice with no AI software. Future studies to identify number of CT referrals made after CXR review with and without AI software assistance, stratified and reported by clinical pathway for both symptomatic (suspected lung cancer) and incidental (no lung cancer suspected) primary care population are required. Ideally these would be of prospective study design, but use of

hospital reported data could be used to retrieve this information retrospectively in the incidental primary care population.

- Number of people referred for follow-up CXR

Two studies[34, 35] reported information on follow up CXR following initial CXR results. In the study by Woznitza 2018[35] all CXRs reported as showing pneumonia had a follow-up CXR suggested in 4-6 weeks to ensure resolution (17/522, 3%). Where a follow-up CXR was suggested, four (22%) were performed with mean time from initial to follow-up CXR of 33.8 days (range 10-49 days). In the 13 other cases follow up was not done at the same institution and authors assumed they had not been undertaken as no reminders were sent.

Bradley 2021[34] reported follow up CXRs performed based on result codes of 2129 initial GP-requested CXRs. Of the 376 patients who had an initial 'negative' result (codes 3 and 4), 98 (26.1%) had at least one further CXR. 370 patients with an initial abnormal finding where non-urgent further review or investigation was advised (code 2) of which 191 (56.1%) had a second CXR. The median duration to second CXR was 42 days (IQR, 28-57). 324 (15.2%) patients across all CXR result codes (1-4) had at least 2 CXRs before diagnosis.[34]

These studies[34, 35] are informative of CXR resource use across multiple pathways, which is useful to consider in future modelling. Whilst Woznitza 2018[35] only had a relatively small sample size in their feasibility study, it was a prospective design and for this measure reported specifically for those on a clinical pathway following 'abnormal' (other diagnosis) CXR results. This would be pathway B (see section 5.3) in the conceptual model. It illustrates that the number of people referred for follow-up CXR does not necessarily equate to resource use, as patient uptake rate is also a factor.

- Number of cancers missed/detected

One (ineligible) study[22] from the clinical effectiveness review reported mean number of cancers detected and found no significant differences with and without AI software (54 cancers, 95% CI 42 to 59; and 46 cancers, 95% CI 38 to 51, respectively).

Of the 1,687 CXR referrals in the study by Woznitza 2018,[35] 17 patients were missed who were subsequently diagnosed with lung cancer. 15 were given normal CXR results and two abnormal (alternative diagnosis) results.

Out of 8682 CXR referrals in the Woznitza 2022[36] study, 48 of the 49 lung cancers diagnosed were detected. The single case that was missed was diagnosed on a subsequent emergency attendance for upper limb deep vein thrombosis.[36]

Foley 2021[33] reported the number of cancers diagnosed by CX code: CX1, 10/1,056 (1%); CX2, 29/288 (10%); and CX3 49/72 (68%). Ten people with lung cancer were given false negative 'normal' results but still referred for CT and received diagnosis. Data on the other 949 patients with negative CXR results who were not referred for CT would be informative (although difficult to obtain) to give total number of false negative results by CX1 code for use in modelling. Similar information would be required for the CX2 result patients.

Future studies with extended follow up and use of patient-level hospital reported data linked to cancer registries would facilitate access and reporting of this information. This may also provide data on stage at diagnosis.

Numbers of false negatives (i.e. lung cancers missed) and stage at diagnosis may both be important outcome measures for use in evidence linkage. This could be used in modelling to inform any association between time to diagnosis and stage shift and to assign appropriate costs and quality of life outcomes by stage at diagnosis.

- Stage of cancer at detection

Bradley 2021[34] found 1,490 (70%) of the 2,129 patients in their study were diagnosed with lung cancer at stage III/IV. Across the four CXR codes used to stratify results of initial CXR these were reported as 1., 981 (70.9%); 2., 259 (70%); 3., 147 (63.9%); and 4., 103 (70.5%). There was no evidence of a statistically significant association between CXR result and stage at diagnosis.[34]

Foley 2021[33] also found no statistical difference between CXR result and stage at diagnosis. Those with advanced stage (IIIc/IV) at diagnosis were reported as CX1, 5 (50%); CX2, 11 (38%); and CX3, 28 (57%) ($p = 0.26$). This was a much smaller sample size than the Bradley 2021[34] study, and advanced stage was defined as IIIc/IV[33] rather than III/IV.[34]

Findings from both studies[33, 34] showed a majority of patients with a normal or abnormal CXR results have advanced stage disease at diagnosis.

- Time to CXR report

Time to CXR report was highly dependent on Trust and service provided. Most had same-day reporting facility for GP requested plain CXR films. Woznitza 2022[36] reported on the RadioX trial which compared immediate reporting and standard reporting to find median report time to CXR report (termed as turnaround time (TAT) in this paper)

This may be a more informative measure than TAT per scan as this impacts more directly on speed at which CT scan is requested.

- Time to CT scan

Time from CXR to CT scan was reported in two studies retrieved in the cost-effectiveness search.[33, 35] Foley 2021[33] found a significant difference in number of days from CXR to CT by CX result code. A mean of 34.6 days for those with CX1, normal but a CT scan is recommended to exclude malignancy; 19.6 days for CX2, alternative diagnosis; and 1.9 days for CX3, suspicious for cancer, was reported.

In contrast, the feasibility study by Woznitza 2018[35] looked at time from CXR to CT scan by reporting strategy for those with a CXR results suspicious for lung cancer. Those who had immediate reporting of their CXR image had a mean of 0.9 days (n=22 mean 0.9 days SD=2.3) until CT scan compared to routine reporting (mean 10.6 days; SD=4.5; p>0.0001).

Whilst these cannot be directly compared, the results of Woznitza 2018[35] are for the equivalent result population of the CX3 in the Foley 2021[33] study. This shows significant variation in time from CXR to CT scan due to department reporting practices alone. In the Foley 2021[33] study there were GP reporting sessions for consultants on most days (Jonathan Rodrigues, Consultant Radiologist, Royal United Hospitals Bath NHS Foundation Trust, personal communication 13.12.2022), suggesting this was more in line with the standard reporting process in the study by Woznitza 2018.[35] However, many other procedural variables between the two different radiology departments are likely to have an impact on these times.

This highlights the need for the real-world clinical context to be taken into consideration in the generation of future evidence to inform these measures. This is relevant for studies of outcomes after CXR both with and without AI, as there is only limited data even in current

practice, which is difficult to generalise due to variation both within and between NHS Trusts.

Results from Foley 2021[33] are useful for modelling purposes as they establish a difference in time between three diverging clinical pathways, up to the point of confirmatory testing by CT scan. This may support evidence linkage to outcomes further in the lung cancer management pathway.

- Time to diagnosis

Four studies[33, 34, 35, 36] from the cost-effectiveness review reported time to diagnosis. In three studies[33, 34, 36] this was calculated as date of initial CXR to date the diagnosis was confirmed (either the date of the diagnostic test or the date on which a clinical diagnosis was confirmed by the lung cancer MDT if no pathological sample was taken). In the smallest of the studies, Woznitza 2018[35] used date of radiological diagnosis confirmed at MDT. Results of histological diagnosis was reported separately, but no data on timing for this was provided.

Foley 2021[33] found a significant difference in time to diagnosis between CX codes, with a mean of 89.7 days for those with CX1, normal but a CT scan is recommended to exclude malignancy; 65.3 days for CX2, alternative diagnosis; and 30.2 days for CX3, suspicious for cancer.

Bradley 2021[34] also reported time to diagnosis by initial CX codes, but used median number of days. For those with code 1. Suspicion of lung cancer identified/urgent investigation needed, 36 days (IQR, 23-63); 2. Abnormality identified/non-urgent investigation indicated including diagnoses of pneumonia or consolidation even if repeat imaging was not explicitly suggested, 93 days (55-154); 3. Abnormality identified but no further investigation/assessment indicated, 211 days (181-296); and 4. Normal CXR, no abnormalities identified 193 days (87-279. When calculated by author defined ‘positive’ (codes 1 and 2) and ‘negative’ (codes 3 and 4) time to diagnosis was 43 days (27-78) and 204 days (105-287), respectively.[34]

Woznitza 2022[36] presented both mean and median days to diagnosis for those who had immediate reporting (IR) of their CXR image and those who had standard reporting (SR).

Mean days for IR were 47.2 (S.D., 35.8) and 81.6 (78.5) for SR. When median days to diagnosis of 32 (IQR, 19-70) for IR and 63 (29-78) were analysed, statistical significance was shown (p=0.03).[36]

Woznitza 2018[35] also looked at mean time to diagnosis for IR and SR, with study findings of 4.1 and 10.6 days, respectively. However, this was for a small sample of 11 patients, and as discussed this was for radiological diagnosis at MDT only, so did not account for additional waiting time due to biopsy.[35]

All four studies[33, 34, 35, 36] reported substantial variation in time to diagnosis, demonstrating this outcome measure can be affected by multiple factors including CXR result and the subsequent diagnostic pathway followed[33, 34] and different reporting practices in a radiology department.[35, 36] Establishing that AI software has an impact on time to diagnosis beyond fluctuating departmental factors, the mechanism by which that impact is produced (through increased test accuracy, reducing report turnaround time or by other means) and quantifying the impact would require a prospective study, in real world settings, ideally across multiple-sites in the UK.

Once established, change in time to diagnosis may support evidence linkage to outcomes further in the lung cancer management pathway.

- Ease of use/acceptability of the software by clinicians

In the UK study[22] ten of the 11 clinicians responded to questions about acceptability of the AI. Eighty percent stated that reporting was not slower when using AI and 90% stated that the AI 'heatmaps' produced were '*helpful to understand the algorithm's attention points*'.

Clinical outcomes for consideration may include:

- Morbidity
- Mortality

Costs will be considered from an NHS and Personal Social Services perspective.

Costs for consideration may include:

- Cost for each AI software available for this indication
- Costs for training staff to use software

- Costs associated with healthcare professional time to read and report CXR
- Costs for diagnostic testing and treatment

Sources for cost and resource use inputs is discussed under Question 5. sub-section.

Summary

Importantly, the EAG did not identify any studies in the clinical effectiveness review that met the inclusion criteria and addressed the outcomes for discussion, highlighting the gap in evidence for all measures of AI software to inform cost-effectiveness analysis at this time.

Four studies[33, 34, 35, 36] from the cost-effectiveness review looking at CXR referral from primary care referrals without AI software informing model parameters, but all had limitations on their applicability due to study type and reported outcomes.

Use of hospital reported data to conduct retrospective studies shows promise to provide good quality information on outcomes under current CXR review practices without AI software. Reporting consistently defined, key clinical pathway outcomes by standardised CXR report codes for would allow comparison between studies and provide more straight forward translation for use in cost-effectiveness modelling.

Coordinated research efforts are required to generate research on all outcome measures identified in for inclusion in the conceptual model. Evidence needs to demonstrate impact on intermediate outcomes over a sustained period of time in the NHS environment to account for differences in outcomes due to the widespread variation in current practices and pathways between individual hospitals sites and Trusts. This can be achieved through well-designed studies, with large sample sizes, conducted over a sufficient period to capture the main outcomes of interest. This would reduce the reliance of evidence linkage which remains particularly weak with regards to impact on stage at diagnosis.

Question 4.

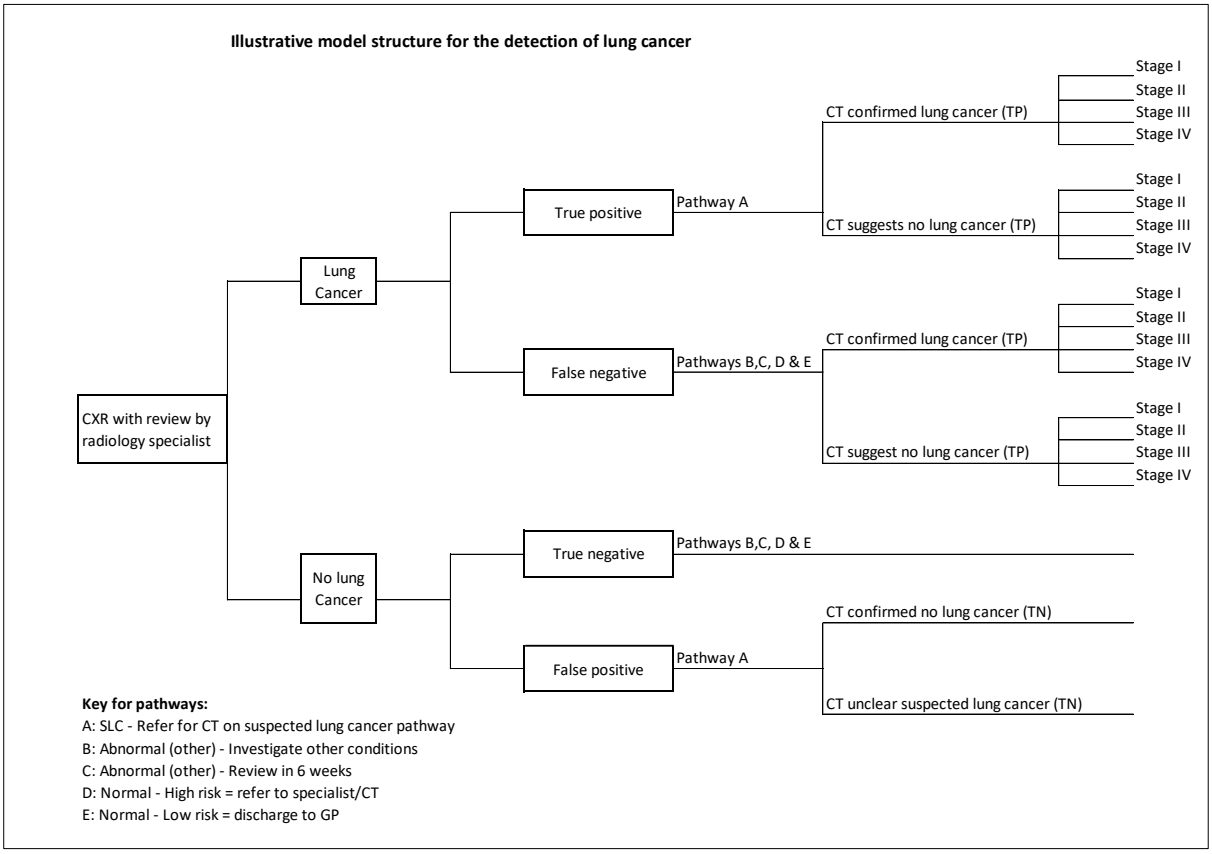
What would a health economic model to estimate the cost-effectiveness of adjunct AI to detect lung cancer look like?

This section describes a conceptual model developed by the EAG to identify the structure and components required in any future health economic models estimating cost-

effectiveness of adjunct AI compared to radiologist or reporting radiographer review alone of CXR images to detect lung cancer.

The proposed structure is suitable for both symptomatic and incidental primary care populations referred by their GP for CXR, with certain model parameters varying where appropriate for the specific population.

The conceptual model follows the illustrative pathways shown below in Figure **4**.



The illustrative pathway for CXR review by radiology specialist with adjunct AI software is identical to the structure presented here for CXR review by radiology specialist alone. If AI was used for triage, an additional step prior to the CXR could be included.

Figure 4: Illustrative model structure for the detection of lung cancer

Strategies

For people undergoing a CXR, the CXR image is read by either a radiology specialist alone (current usual practice) or radiology specialist with adjunct AI software.

Proposed model structure

A decision tree structure is used to depict the pathway from CXR imaging and review, to point of diagnosis. We considered a decision tree structure appropriate to capture the short-term costs and benefits associated with the strategies to identify people with lung cancer.

A positive CXR result (findings suspicious of lung cancer) follows pathway A, where a CT scan confirms the positive result and provides provisional staging. A utility decrement is applied to a positive result lasting until treatment. Treatment according to stage at diagnosis then commences when utility values for that stage are attributed for true positive cases. False positive cases revert back to general population utility values.

People with false negative results follow pathways B,C, D or E depending on whether findings are reported as 'normal' or 'abnormal (alternative diagnosis)'. These people eventually undergo a CT scan either as part of further clinical investigations along their respective pathways, or they are assumed to present at an emergency department later. Any false negatives not detected at first CT scan along any pathway are also assumed to present later as an emergency. These pathways are longer than the most direct route to diagnosis (pathway A) and it is assumed that the delay in time-to-diagnosis confers a stage-shift for a proportion of these people. Treatment then commences by stage at diagnosis.

People who receive a false positive result at CXR imaging also follow pathway A and go on to receive a CT scan as a minimum further investigation, with a proportion who undergo further testing (e.g., PET-scan, biopsy, bronchoscopy) until a true negative lung cancer diagnosis is confirmed. A temporary utility decrement is applied for a false positive test result for the duration until a confirmatory test is received showing no lung cancer present. A utility decrement associated with further invasive diagnostic procedures (biopsy and bronchoscopy) is applied to people with true positive results and a proportion of those with false positive results.

As for those people with false negative results, people with true negative results follow pathways B,C, D or E depending on whether findings are reported as 'normal' or 'abnormal (alternative diagnosis)'. Additional testing is specific to each pathway.

Pathways A, B, C, D and E are described in detail in section 5.3. Within the model, separate costs and health related quality of life (HRQoL) outcomes are assigned to each pathway. All pathways which lead to a diagnosis of lung cancer complete the decision tree at a fast-track lung cancer clinic. Total costs and HRQoL outcomes (expressed as quality-adjusted life years (QALYs)) to point of diagnosis are accrued according to the proportion of people assigned to each pathway as a result of CXR review by the two strategies under comparison.

At the end of the decision tree branches, long-term treatment costs and utility values over a five-year time horizon are assigned based on stage of lung cancer at diagnosis. These are added to those accumulated during the diagnostic component of the model to provide overall outcomes for each strategy.

Results of subsequent analysis (in a fully executable model) would be presented in terms of an incremental cost-effectiveness ratio (ICER), where the difference between total costs of CXR review by radiology specialists with and without adjunct AI, is divided by the difference between total QALYs for each, to give a cost per QALY figure. Prices would be based on the current cost year with discounting of cost and outcomes applied at 3.5% over the total model time horizon in line with NICE reference case.[18]

For the conceptual model information is required about the prevalence of lung cancer and the performance of radiology specialists to detect findings indicative of lung cancer on review of CXRs both with and without AI used as an adjunct. These inputs are specific to the population of interest, so figures are required for prevalence and diagnostic accuracy in both symptomatic and incidental primary care populations.

Prevalence figures used in the literature are sourced by Bajre 2017[30] for use in their economic evaluation from Field 2016,[46] and by Geppert 2022[45] for modelling in the DAP060 AI for chest CT diagnostic assessment review from Horeweg 2014.[47] Both sources[46, 47] contain estimates of prevalence for lung cancer in the screening population. For modelling purposes in Bajre 2017[30] and Geppert 2022,[45] these prevalence estimates are assumed to be the same for their population of interest. The EAG did not find any more relevant sources, but searches were not exhaustive and more recent estimates of prevalence in the UK population would be advisable for use in future modelling.

For the specific clinical pathways (A,B,C,D & E) people may follow through the decision tree, information on costs and resource use of diagnostic tests and clinical management input is required. The proportion of people taking each pathway and the mean time from initial CXR to diagnosis is also required for each of these pathways, under each strategy.

An example using clinical pathway A

CXR findings suspected of malignancy are flagged to secondary care lung cancer teams who request the CT scan and await referral to the suspected lung cancer clinic from the GP. Once reported CT scans are triaged by lung cancer team consultants. If they suggest probable lung cancer, an urgent lung cancer team appointment is arranged with appropriate tests e.g. lung function tests, planned biopsy (EBUS) (Alberto Alonso, Consultant Radiologist, Manchester Hospital NHS Foundation Trust, personal communication 19.12.2022). Diagnosis, histopathological staging and treatment options are then discussed at the fast-track lung cancer MDT clinic.[16]

This process incurs the cost per person of a CT scan (£153), lung function tests (£285) and biopsy (£1670).[48]

Input to direct these further tests is required by the secondary care lung cancer team on two occasions; 1) to review CXR results, refer for CT scan and notify GP to make suspected lung cancer pathways referral and 2) to review CT scan results and refer for lung function tests and biopsy in prior to fast-track lung cancer MDT clinic review. The unit cost of a lung cancer MDT meeting (£146)[48] or part thereof, would be assigned for both encounters with the lung cancer team and the fast-track clinic team. Average times to discuss a case during these meetings is necessary for more accurate costing.

Similarly, utility decrements are also assigned to pathway A. Suspicious lung cancer findings on CXR attract a disutility of –0.063[49] applied over the length of time until confirmatory diagnosis. A disutility of -0.2 is applied for biopsy investigation for a period of three months.[50, 51]

The total costs and QALYs accrued are then attributed to the proportion of people in the model who take pathway A as a true positive or false positive case.

Table 8: Input parameters required to populate conceptual model

	Population			
	Primary Care Symptomatic		**Primary Care Incidental**	
Parameter	Value	Source	Value	Source
Prevalence				
Radiology review without AI software				
Sensitivity				
Specificity				
Radiology review with AI software				
Sensitivity				
Specificity				
Proportion of people following each pathway after CXR review without AI software				
Pathway A				
Pathway B				
Pathway C				
Pathway D				
Pathway E				
Proportion of people following each pathway after CXR review with AI software				
Pathway A				
Pathway B				
Pathway C				
Pathway D				
Pathway E				
Proportion of lung cancers diagnosed at stages I, II, III and IV after CXR review without AI software				
	Stage I	II	III	IV
Pathway A				
Pathway B				
Pathway C				
Pathway D				
Pathway E				
Proportion of lung cancers diagnosed at stages I, II, III and IV after CXR review with AI software				
	Stage I	II	III	IV
Pathway A				
Pathway B				
Pathway C				
Pathway D				
Pathway E				
Utility values for lung cancer diagnosed at stages I, II, III and IV				
Utility value	Stage I	II	III	IV

The conceptual model presented captures the following important outcomes in the diagnostic process.

Clinical outputs from the model

- Number of false positives
- Number of additional CT scans
- Number of people referred for follow-up CXR
- Number of people identified as 'normal' (no lung cancer present) and discharged
- Number of cancers missed and detected
- Proportion of cancers detected at each stage

Long-term outcomes from the model

- Total costs per strategy
- Total QALYs per strategy
- Costs per QALY

These outcomes would be based on a cohort of 1000 patients entering the model.

Question 5.

What are the cost and resource use considerations relating to use of adjunct AI to detect lung cancer?

This section identifies the costs and resource use of adding AI software to CXR review taking an NHS and Personal Social Services perspective. Costs are required for each AI software, costs for training staff to use software, resource use and costs associated with healthcare professional time to read and report CXR, and costs for diagnostic testing and treatment. All costs are presented in 2021 prices. Costs obtained from the literature were uprated to current prices using the Hospital and Community Health Services index from Unit Costs of Health and Social Care 2021.[52] Cost categories are listed with resource use considerations discussed alongside and any potential sources of information identified.

Cost of software

AI software costs were obtained directly from the companies. Five of the 14 companies identified in the Final Scope[10] registered as stakeholders in this EVA and provided cost information to the EAG via NICE communications (Annalise AI, Behold AI, Infervision, Lunit Inc., and Siemens Healthineers).

Pricing structures were either fixed annual subscription fees (Annalise AI, Behold AI, Infervision and Siemens Healthineers) or volume-based annual pricing tiers (Infervision and Lunit Inc.). All companies charge a one-off implementation fee in the first year which covers installation, integration to existing PACS/RIS, and staff training. On-going subscription costs are renewable on an annual basis, with fees covering software licensing, annual maintenance, support and updates. Pricing is calculated per Trust by Annalise AI, Infervision, Lunit Inc.and Siemens Healthineers. In contrast Behold AI's implementation and subscription fees are per hospital, with a 30,000 annual CXR volume allocation.

[redacted]

Annual subscription cost is dependent on the volume of CXRs to be processed in either each Trust (Annalise AI, Infervision, Lunit Inc.and Siemens Healthcare) or each hospital (Behold AI) annually. The resource use would, therefore, be determined by the number of primary care referrals for CXR for the symptomatic and incidental populations per year.

This information is available through Trust databases and has been reported in the literature through retrospective database studies.[33, 34]

Table 9 shows disaggregated costs of AI software by company based on a volume of 25,000 CXR images per NHS Trust.

Table 9: Costs of AI software by company based on a volume of 25,000 CXR images per NHS Trust

Company, Technology name, (Tech use)	**One-off set up cost/ implementation fee**	**Annual subscription (based on volume 25k images)**	**Cost per exam**	**Total first year cost**	**Indicative cost per image (non-discounted) (5yr average)**
Annalise AI, Annalise Enterprise & Triage, (CADe & CAST)	£5,000 - £25,000	£51,250*	N/A	£66,250	£2.17
Behold.ai, Red dot, (CADe & CAST)	£10,000**	£60,000**	N/A	£70,000	£2.48
Infervision InferRead DR, (CADe)	£3,000	£16,000 (license fee) £6,000 (maintenance fee)	N/A £1.00	£25,000 £34,000#	£0.90 £1.24
Lunit Inc. Lunit, INSIGHT CXR, (CADe)	£6,000	£16,750 to £50,000	£0.67 to £2	£22,750 to £56,000	£0.72 to £2.05
Siemens Healthcare, AI-RAD, (CADx)	£2,400	£12,000*	N/A	£14,400	£0.50

*Based on tier pricing of 'up to' 25,000 images per year

**Per hospital cost

Cost per image price structure (includes one off implementation fee and annual maintenance fee)

Cost of staff training

Staff training is provided by the AI software companies and the cost included in the one-off implementation fee (see section X). Companies reported that training time for radiologists/ reporting radiographers was 1-hour for Lunit, and 30 minutes for Infervision. For Behold AI, no training time was given. Instead they advised a training deck is customised for each Trust, used to train designated trainers from each organisation then the deck given to the trainers to provide training to their respective radiologists.

Under the assumption that training is undertaken during protected staff-training time within radiology departments, no further costs would be attributed beyond that of the implementation fee.

Cost of staff time to read and report CXR

The hourly cost of a radiologist or reporting radiographer was obtained from the literature. Two methods were identified which had been used in previous economic evaluations. [30, 53] In the first of these, Bajre and colleagues[30] used the figure of £156 per hour for a radiologist and £53 for a band 7 reporting radiographer. This was originally calculated by Lockwood[53] based on salary, on-costs and education for the 2015/16 cost year. In the second economic evaluation, the hourly cost of a band 9 radiographer (£147) from the Personal Social Services Research Unit (PSSRU) Unit costs of health and social care 2021[54] was used as a proxy for a radiologist.[45]

Cost of staff time to read and report a single CXR can then calculated using turnaround time (TAT). Published evidence has suggested no statistically significant difference in reading times of chest x-rays between readers with and without AI.[24, 25] However, these are data are from two studies[24, 25] which do not meet the inclusion criteria for the present EVA, and are of uncertain applicability to clinical practice (see section 4.4).
Feedback from Diagnostics Advisory Committee (DAC) specialist members suggested timings without the use of AI of between one minute on average, faster for normal and slower for very abnormal, up to five minutes. From the literature, Bajre et al.[30] assumed a two-minute reporting time for both radiologists and reporting radiographers.

Cost of further diagnostic tests

Following the initial CXR, further testing may be required. This could include additional CXR, CT-scan of chest, CT-scan of abdomen (performed with or without contrast), PET-scan, bronchoscopy, and biopsy with various combinations of each possible.

To direct these further tests, clinical input from the GP, respiratory specialists, radiologists, and appropriate multidisciplinary teams (MDTs) are required. Costs of these services can be obtained from the National schedule of NHS costs 2020/21[48] and the PSSRU Unit costs of health and social care 2021.[54]

Costs of further tests are dependent upon outcomes along the clinical pathway including the:

- number of people referred for a CT scan
- number of people referred for follow-up CXR
- number of people identified as 'normal' (no lung cancer present) and discharged
- number of cancers missed/detected
- stage of cancer at detection

No evidence was identified in the clinical search which addressed these outcomes as a result of AI software assistance in the reading of CXRs. We therefore have no evidence on which to determine whether the use of adjunct AI will increase, decrease, or not affect the number of people requiring additional testing.

Cost of treatment (including costs of any adverse events)

Total treatment costs are assigned according to stage of disease.

Several sources were identified in the literature. Bajre et al.[30] and Geppert 2022[45] used Cancer Research UK 2014[55] values originally reported in the 2014/15 price year and included cost of retreatment.

Snowsill and colleagues[31] used figures based on a two-year costing approach, with index year costs from a UK teaching hospital[56] and second year costs estimated from the index

year using a subsequent year ratio from database analysis in England.[57] The same authors in an interim update to the UK National Screening Committee (NSC)[32] also used a five-year micro-costing approach with resource use based on the most recent NLCA secondary care estimates 55-75 year age range to reflect more modern available treatment options including immunotherapies.

See Table 10 for a summary of costs required for the proposed model.

Table 10: Costs required for the proposed model

Parameter	Value	Source
Healthcare professional		
GP consultation	£39	PSSRU 2021/22[54] (per patient contact of 9.22 mins)
Radiologist consultation	£147	PSSRU 2021/22[54] (cost per working hour (£147) for a Band 9 radiographer as a proxy for a radiologist)
Multidisciplinary team	£146	National schedule of NHS costs 20/21[48] (CDMT_OTH other cancer MDT meetings)
Other tests		
X-ray	£45	NHS reference schedule 20/21[48] (Direct access plain film)
CT scan (single area, with Contrast)	£153	NHS reference schedule 20/21[48] (RD21A- computerised tomography scan of one area, with post-contrast, 19 years and over
Computerised Tomography Scan of two Areas, without Contrast	£127	NHS reference schedule 20/21[48] (RD23Z- computerised tomography scan of two areas, without contrast)
Computerised Tomography Scan of two Areas, with Contrast	£153	NHS reference schedule 20/21[48] (RD24Z- computerised tomography scan of two areas, with contrast)
Guided needle biopsy	£1670	NHS reference schedule 20/21[48] (DZ71Z- minor thoracic procedure, guided needle biopsy)
Lung Function Tests	£285	NHS reference schedule 20/21[48] (DZ52Z – Full lung function testing)
Bronchoscopy	£1679	NHS reference schedule

		20/21[48] (DZ70Z- Endobronchial ultrasound examination of mediastinum)
PET scan	£1161	NHS reference schedule 20/21[48] RN01a- PET-CT of one area, 19 years and over
Treatment		
Stage I	£20,928	UK NSC external review: interim report[32]
Stage II	£29,757	
Stage III	£32,830	
Stage IV	£21,838	
CT, computed tomography; NSC, National Screening Committee; PET-CT, positron emission tomography and computed tomography; PSSRU, Personal Social Services Research Unit		

For use in any future modelling, all sources where costs are obtained from the literature will require uprating to current prices at the time using the Hospital and Community Health Services (HCHS) index from the most recent publication of the Unit Costs of Health and Social Care.

Summary

Potential sources to inform all unit costs for the cost parameters in the conceptual model proposed by the EAG have been identified. Primarily, these costs can be obtained from the literature, published national index costs and directly from AI software companies.

Evidence to support resource use relating to adjunct AI to detect lung cancer was not identified. Therefore, total values of cost inputs for all cost parameters could not be calculated.

No evidence was found to determine what, if any, effect AI will have on resource use and in what direction this might take with respect to costs. At this stage all we able to determine is that AI represents a new cost, as AI software needs to be purchased and used an addition to the costs and resources consumed in the current clinical pathway.

5.5. Results of potential budget impact assessment

Five companies provided cost information to NICE as part of the DAP request for information process, all of whom responded to the EAG's clarifying questions (Annalise AI, Behold AI, Infervision, Lunit Inc., and Siemens Healthineers). This provided more certainty in the calculation of costs of these technologies performed by the EAG. In total there was sufficient information for six different price estimates (Infervision provided two different pricing structure options).

AI software is intended as an adjunct to the existing CXR review process conducted by a qualified radiology specialist. The ultimate diagnostic decision is made by the radiology specialist, the cost of which is a assumed to be constant in both current and future practice if AI software were to be implemented. This assumption was made as no evidence was found in our review of any change in resource use due to AI software. With this the case, only the additional costs of AI software are considered here.

As discussed in the conceptual modelling process results, there was no available evidence found to inform any changes to progression through the clinical pathway due to the intervention in this population. Therefore, onward health-related service use, diagnostic and treatment costs are assumed to stay the same for the purposes of this analysis. However, for the purposes of any future modelling, costs that may need to be considered include cost of CT scans, CT surveillance for lung nodules detected, cost of further invasive tests e.g. biopsy, and treatment for different stages of lung cancer at diagnosis.

Change in test accuracy may result in increased sensitivity with AI software assistance and potentially identify more cancers/nodules, or decreased specificity (i.e. because of an increase in false positives) wherein more people could be referred for a CT scan with an associated cost implication.

Several studies were retrieved during the literature search which appeared to provide sufficient data on which budget impact at an individual institution level could be calculated.[33-36] These studies have previously been summarised. The study by Foley and colleagues 2021[33] was chosen to base the budget impact case on as the trust wide annual referral number for the appropriate populations (both suspected lung cancer and incidental primary care) were clearly provided. It was also not restricted to those who had a confirmed

diagnosis of lung cancer, as with the Bradley et al. (2021)[34] study, and authors responded to clarifying questions from the EAG upon contact to ensure greater accuracy in interpretation of study results.

Foley and colleagues[33] conducted a retrospective review of 16,945 CXRs referred from primary care and performed across all sites at the Royal United Hospitals Bath NHS Foundation Trust, between June 1, 2018, and May 31, 2019. 1,488 of these were referred for suspected lung cancer.

Upon contact with corresponding authors, annual GP referral data for CXR to the Royal United Hospitals Bath NHS Foundation Trust, including break down of those referred for suspected lung cancer, was provided to the EAG for the period January to December, 2019 to 2022, inclusive (Richard Wood, PACS Manager, Royal United Hospitals Bath NHS Foundation Trust, personal communication 6.02.2023). The EAG intended to calculate budget impact estimates based on these exact numbers, including those reported in the study[33] for 2018, as a example of the first five years of AI software implementation at single NHS Trust. However, due to substantial variation in numbers referred over this time as services were impacted by the Covid-19 pandemic, the EAG decided to use a conservative assumption where the annual referral number from primary care was kept constant at 16,945 over the five years for this analysis.

Results are presented in Table 11, with anticipated budget impact at NHS Trust level for both symptomatic (suspected lung cancer) and incidental primary care population CXR referrals shown in the final column.

Table 11: Anticipated budget impact of AI software at NHS Trust level for all GP-referred CXR

Company **Technology name** **(Tech use)**	**One-off set up cost/ implementation fee**	**Annual subscription (based on volume 16,945 images)**	**Cost per exam**	**Total first year cost** **[VAT applied at 20%]**	**Cost over first 5 years (non-discounted) (based on volume 16,945 images per yr)** **[VAT applied at 20%]**

Annalise AI **Annalise Enterprise & Triage** **(CADe & CAST)**	£5,000 - £25,000	£51,250*	N/A	£66,250 (assuming mean implementation fee) [£79,500]	£271,250 [£325,500]
Behold.ai Red dot (CADe & CAST)	£10,000	£60,000	N/A	£70,000 [£84,000]	£310,000 [£372,000]
Infervision InferRead DR **(CADe)**	£3,000	£16,000 (license fee) £6,000 (maintenance fee)	N/A £1.50	£25,000 [£30,000] £34,418 [£41,302]	£113,000 [£135,600] £160,088 [£192,106]
Lunit Inc. Lunit INSIGHT CXR **(CADe)****	£6,000	£11,353 to £33,890	£0.67 to £2	£17,353 to £39,890 [£20,824 to £47,868]	£62,766 to £175,450 [£75,319 to £210,540]
Siemens Healthcare **AI-RAD (CADx)**	£2,400	£12,000*	N/A	£14,400 [£17,280]	£62,400 [£74,880]

*Based on tier pricing of 'up to' 25,000 images per year; ** reports minimum to maximum

Results are presented in Table 12, with anticipated budget impact over the first five years at NHS Trust level separately for symptomatic, incidental and total primary care population CXR referrals.

Table 12: Anticipated budget impact of AI software at NHS Trust level for symptomatic, incidental and whole population GP-referred CXR

Company **Technology name** **(Tech use)**	**Cost over first 5 years for symptomatic primary care population**	**Cost over first 5 years for incidental primary care population**	**Cost over first 5 years for all primary care population referrals**
Annalise AI **Annalise Enterprise & Triage** **(CADe & CAST)**	NDA	£325,500	£325,500
Behold.ai Red dot **(CADe & CAST)**	£372,000	£372,000	£372,000
Infervision **InferRead DR** **(CADe)**	£135,600 £52,992	£135,600 £178,713	£135,600 £192,105
Lunit Inc. Lunit INSIGHT CXR **(CADe)****	£13,182 to £25,056	£69,337 to £192,684	£75,319 to £210,540
Siemens Healthcare **AI-RAD (CADx)**	£26,880	£74,880	£74,880
NDA = No data available. Non-discounted costs, VAT included at rate of 20%. Total population n = 16,945, symptomatic population n = 1,488, incidental population n = 15,457. **report minimum to maximum			

The budget impact estimates for the whole primary care population referred for CXR is not expected to total the sum of the symptomatic and incidental populations (see Table 12). This is due to the assumption during calculation that AI software is only approved for use in that specific sub-population. Use of volume-based pricing structures also means that the cost of AI software implementation and use over five years would be the same with [redacted] for each of the symptomatic and incidental populations alone as it would be for the whole primary care population.

Summary

Budget impact results vary greatly between companies, but the EAG cautions against direct comparison, as the AI software presented has varying capabilities and some may be used in different positions early in the diagnostic pathway. For example, Sieman's AI software points to a region of interest on the CXR whereas Annalise AI software identifies a specific location, gives characteristics of the anomaly on CXR and provides a preliminary diagnosis and rating of confidence when used in concurrent review of images. Similarly, Behold AI and Annalise AI software can provide triage of CXR images prior to radiology specialist review in order to prioritise reporting, as well as assist with detection of abnormalities and diagnosis. These differing capabilities may affect the way the AI software is used in practice with a variety of practical, clinical and cost implications later in the diagnostic pathway. Therefore, without future modelling it is unclear how budget impact estimates for different AI software brands might be comparable.

6. Discussion

6.1. Statement of principal findings

6.1.1. Test accuracy, practical implications, and clinical effectiveness

No studies met the systematic review inclusion criteria. There is currently no evidence on the use of adjunct AI software for the detection of suspected lung cancer on CXR in either people referred from primary care with symptoms of lung cancer or in people referred from primary care for other reasons. This finding, however, satisfies the secondary aim of this review, which was to identify evidence gaps in this field and inform future research. This is discussed in more detail below.

To provide context to the decision problem, summary results were presented from six studies that did not meet the review inclusion criteria due to unclear populations but were selected for discussion post hoc. The referral status and symptom status of the study participants are unknown, but the studies did provide comparisons of CXRs read by radiologists with and without the use of commercial AI software. Few outcomes were reported in these studies. They provide some insight into two of the key questions of this EVA:

- Question 1, what is the test accuracy and test failure rate of adjunct AI software to detect lung on CXR?
- Question 2, what are the practical implications of adjunct AI software to detect lung cancer on CXR, respectively?

None of the studies provided evidence on the clinical effectiveness of adjunct AI software applied to CXR (Question 3).

For question 1, one study reported a higher sensitivity for lung cancer detection by readers with adjunct AI compared to readers alone, with no difference in specificity or cancer detection rate.[22] No significant between-group differences were found in test accuracy metrics in the five studies assessing lung nodules.[24, 25] [23] [26] (Siemens, 2022)

For question 2, no significant between-group differences were found for reading time, [24, 25] or hypothetical referrals for CT scan.[22, 24] Data from one study indicated that clinicians generally responded positively to the use of AI software.[22]

This synopsis of study results is illustrative only of the type of evidence that is currently available on commercial AI to aid the interpretation of CXR. Caution is required in extrapolating from these studies as not only did they not meet the review inclusion criteria, but there were also differences between the studies, and limitations within them. For example, some studies included nodules with differing levels of detection difficulty from easy to challenging, while others excluded images where nodules were below a certain size, all studies used retrospective designs, data were reported from the mean performance across several readers with varying degrees of experience, readers had their findings from the first reading present at the time of the second reading, and there was a lack of detailed reporting of key results. There were also differences in the reference standards, making comparisons between studies difficult. Further, generalisability to the UK primary care referred population is unclear in all six summarised studies, and generalisability to the UK population overall is likely to be low in three studies that were conducted in Korea.

6.1.2. Conceptual cost-effectiveness modelling

The conceptual modelling process aimed to explore both the structure and evidence requirements for parameter inputs for future model development. There was no evidence available on AI software impact on any of the intermediate outcomes identified to inform parameterisation. Results of EAG searches for evidence to inform these outcomes for the comparator alone (i.e., radiology specialist review of CXR in the detection of lung cancer in the primary care population) varied in study design, and the way outcome measures were reported, limiting the way data could be used.

A simplistic model structure was outlined due to the paucity of evidence and tentative links to long term outcomes.

Key points:

- AI needs to show changes to intermediate outcomes over a sustained period of time in the NHS environment as pathway variation and clinical practice/structure in radiology departments varies considerably between Trust and individual site. Unless evidence is produced which is statistically powered to account for a difference in outcomes due to the current variation, evidence linkage to improved outcomes which may demonstrate cost-effectiveness cannot be made. Ideally this would be in

the form of a large scale, multi-site, UK based clinical trial with AI software as an adjunct to radiology specialist review compared directly with existing practice.

- It is not clear that evidence to suggest stage shift in detection of lung cancer can be achieved through CXR identification of suspected lung cancer in any event.

6.2. Strengths and limitations

6.2.1. Strengths

- Extensive searches including electronic databases, existing reviews, company submissions and known studies which reduces the risk of missing studies.
- Clinical experts were involved in the review and asked to provide details of any potentially eligible studies.

6.2.2. Limitations

- This review employed rapid evidence synthesis methods.[58] While this approach is used internationally by policy-makers to make expedited assessments of evidence,[59, 60] it is not without risks. In the present review, one reviewer conducted all elements of the review in full (i.e., title/abstracting sifting, full text assessment), with a second reviewer assessing/checking 20% of each review task. Therefore, 80% of review tasks were only conducted by a single reviewer. Any errors made by the first reviewer relating to this 80% would not be detected. As such there is the possibility that eligible studies may have been missed.
- We only searched for and included studies published in English language. Therefore, we do not know if there are relevant papers in other languages.
- Targeted searches were used to retrieve a manageable number of records to screen. Therefore, it is possible that some studies (for example, broad reviews) were not retrieved. To counter this, we used different combinations of concepts, sources, and search methods, and tested the overall search strategy's ability to retrieve a set of known studies (found by a variety of methods during the scoping stage).
- Owing to the abridged timescale and limitations in resource for the evidence reviewing processes of this pilot EVA there was no opportunity to follow-up any uncertainties in studies with their authors or to seek further clarification to

responses received from the few companies who provided submissions. Additional time was required to clarify the complex eligibility criteria in the scope before the protocol was signed off and this also impacted on the reviewing timescale.

- As no studies met the eligibility criteria for the review, a pragmatic decision was taken following discussions with NICE to apply additional criteria to the excluded studies to select evidence closest to the review eligibility criteria. This selection process was iterative and involved discussion between two reviewers but was undertaken in the absence of *a priori* defined criteria. As already discussed, studies that were summarised were those where the population referral route and symptom status for the CXR was unknown (not reported). These populations are likely to be no different from other excluded studies with better descriptions of their populations. Also, only summary results were extracted and there was no formal risk of bias tool applied to these studies. These results are illustrative only and results do not provide evidence on the use of adjunct AI software for the detection of suspected lung cancer on CXR in people referred from primary care.
- Selection of cost effectiveness studies was undertaken by one reviewer with wide inclusion/exclusion criteria aimed at pragmatic identification of literature to support development of a conceptual model and inform a rudimentary budget impact analysis of AI software in the NHS, UK. Therefore, it is likely without the rigorously methodology of systematic review processes including quality appraisal. There may be biases from an individual reviewer, and studies identified in this report may not be fully representative of all those available. Through additional searches of references lists of identified studies, publication bibliographies of relevant authorship, several targeted searches and liaison with specialist committee members and clinical experts, the EAG endeavoured to mitigate the risk of missing pertinent evidence for this report.

6.2.3. Limitations of evidence base

This review found no evidence on which to assess artificial intelligence software for analysing CXR to identify suspected lung cancer amongst people referred from primary care.

6.3. Uncertainties

This review aimed to assess the test accuracy and test failure rates of adjunct AI software to detect lung cancer or lung nodules on CXR, the practical implications, the clinical effectiveness of adjunct AI software in people referred from primary care, and to develop a conceptual model. No evidence was found on any of these. Therefore, uncertainties remain regarding all review questions.

The evidence that was summarised to provide some insight into the above was limited to three of the 14 eligible interventions. There was no eligible evidence identified for the following AI software: Annalise CXR (annalise.ai), Auto Lung Nodule Detection (Samsung), ChestLink Radiology Automation (Oxipit), ChestView (GLEAMER), Chest X-ray (Rayscape), ClearRead Xray – Detect (Riverain Technologies), InferRead DR Chest (Infervision), Milvue Suite (Milvue), qXR (Qure.ai), SenseCare-Chest DR Pro (SenseTime), VUNO Med-Chest X-Ray (VUNO).

Resource use associated with progression through clinical pathways was highly uncertain due to lack of evidence and difficult to establish for CXR alone (due to the large number and complexity of clinical pathways possible to diagnosis of lung cancer). Costs for individual elements in the pathway were sourced from published sources used in previous technology assessments, but without robust resource use data this limits certainty in overall cost estimates. Long term treatment costs, calculated by stage at diagnosis, are widely used in the literature with recent updates to these. However, there is only weak and limited evidence to suggest CXR to stage shift at diagnosis.

Due to the lack of evidence for all inputs only a simple conceptual model could be attempted. This by necessity under-estimates the complexity of the pathways and creates uncertainty as to whether this would be the optimum modelling to capture the practical implications question.

6.4. Equality, diversity, and inclusion

We know that an equitable, diverse, and inclusive research group is a more innovative and successful one. Therefore, we integrate ED&I across our workforce, our review products and academic output. We embrace diversity of background, perspective, culture, and

experience, and together with our University and Health and Social Care partners, we work to address inequity.

We provide our team a range of opportunities at different career stages and different levels of commitment; and provide implicit bias training for all team members. We provide flexible research training and opportunities for innovative methodological design work so that everyone can engage in methods development, irrespective of circumstances and career stage. We expect that all line managers and mentors have supervising / mentoring training and can provide confidential and non-judgemental support.

We have built on our strong institutional inclusion and diversity policies to maximise participation of traditionally marginalised groups, and to identify any barriers to develop a supportive culture for new researchers, including encouragement of flexible work arrangements where relevant.

University of Warwick holds silver Athena Swan charter status. Our ED&I policies are regularly reviewed, and awareness is promoted through newsletters and weekly circulars. Warwick Evidence proactively harnesses the research capacity development resources within the university (e.g., mentoring, reverse mentoring, shadowing, strengths profiling) and align these with NIHR Academy systems.

6.5. Patient and public involvement

The short timeline of this EVA meant there was insufficient time to engage patient and public advisors. However, the NICE specialist committee for this assessment included patient representatives who were involved in defining the scope.

7. Conclusions

7.1. Implications for service provision

There is widespread variation in existing service provision both within and across Trusts. Changes in departmental practices alone have been shown to have an impact on outcome measures along the lung cancer diagnostic pathway and have been used positively to try and improve lung cancer diagnosis times.

No evidence was identified in this review to suggest what the impact of AI software as an adjunct to CXR review might have on any stage of the diagnostic pathway.

With a complete lack of evidence on AI software, the impact on service provision is unknown but may have significant implications in terms of progression through diagnostic pathways, resource use, costs, and patient outcomes.

7.2. Suggested research priorities

Given the absence of any eligible evidence on the topic of this EVA, the following research priorities are suggested:

- Assessment of the test accuracy of specialist radiologist with adjunct AI software compared with specialist radiologist without AI software, conducted with participants who reflect those seen in clinical practice. Evidence within these studies should also collected data on the types/characteristics of cancers and nodules that are detected by AI, and the test failure rates of AI. Ideally, this information would come from prospective studies.
- Assessment of the effects that adjunct AI software has on clinical decision-making, and its acceptability to clinicians.
- Assessment of the clinical effectiveness of adjunct AI software to reduce patient mortality and morbidity, and to improve health-related quality of life.
- Large scale studies which support evidence linkage from CXR review to stage at diagnosis, with intermediate outcomes of time to CT scan and time to diagnosis with or without AI software are required. Retrospective audit data from NHS Trusts with sufficient data to link these outcomes in the target population could be undertaken before introduction of AI software. Prospective trials ensuring this data is collected along the full pathway would be the most robust to determine any impact of AI software on these outcomes to account for wide variation in progress through clinical pathways to diagnosis.
- Studies which evaluate quality of life outcomes for people diagnosed with lung cancer by stage of the disease in the UK population are also required.

8. References

1. Cancer Research UK. *Lung cancer*. 2019. URL: https://www.cancerresearchuk.org/about-cancer/lung-cancer (Accessed 14 November 2022).
2. Clinical Knowledge Summaries. *Lung and pleural cancers - recognition and referral*. National Institute for Health and Care Excellence; 2021. URL: https://cks.nice.org.uk/topics/lung-pleural-cancers-recognition-referral/ (Accessed 14 November 2022).
3. NHS website. *Overview: Lung cancer*. 2022. URL: https://www.nhs.uk/conditions/lung-cancer/ (Accessed 14 November 2022).
4. Cancer Research UK. *Lung cancer incidence by age*. 2021. URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer/incidence#heading-One (Accessed 23 November 2022).
5. NHS England. *NHS Long Term Plan*. 2021. URL: https://www.longtermplan.nhs.uk/ (Accessed 14 November 2022).
6. IBM Cloud Education. *Artificial Intelligence (AI)*. 2020. URL: https://www.ibm.com/uk-en/cloud/learn/what-is-artificial-intelligence (Accessed 14 November 2022).
7. Farhat H, Sakr GE, Kilany R. Deep learning applications in pulmonary medical imaging: recent updates and insights on COVID-19. *Mach Vis Appl* 2020;**31**(6):53. http://dx.doi.org/10.1007/s00138-020-01101-5
8. National Institute for Health and Care Excellence. *Lung cancer: diagnosis and management: NICE guideline [NG122]*. 2022. URL: https://www.nice.org.uk/guidance/ng122 (Accessed 14 November 2022).
9. National Institute for Health and Care Excellence. *Suspected cancer: recognition and referral NICE guideline [NG12]*. 2021. URL: https://www.nice.org.uk/guidance/ng12 (Accessed 14 November 2022).
10. National Institute for Health and Care Excellence. *Artificial Intelligence software for analysing chest X-ray images to identify suspected lung cancer: Final scope*. 2022. URL: https://www.nice.org.uk/guidance/gid-dg10065/documents/final-scope (Accessed 23 November 2022).
11. National Institute for Health and Care Excellence, King's Technology Evaluation Centre. *Artificial intelligence for analysing chest X-ray images: Medtech innovation briefing [MIB292]*. National Institute for Health and Care Excellence; 2022. URL: https://www.nice.org.uk/advice/mib292 (Accessed 15 November 2022).
12. Aggarwal R, Sounderajah V, Martin G, Ting DSW, Karthikesalingam A, King D*, et al.* Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. *NPJ Digit Med* 2021;**4**(1):65. http://dx.doi.org/10.1038/s41746-021-00438-z
13. The Joanna Briggs Institute. *JBI Critical appraisal tools*. 2021. URL: https://jbi.global/critical-appraisal-tools (Accessed 14 November 2022).
14. National Institute for Health and Care Excellence. *Early Value Assessment report commissioned by the NIHR Evidence Synthesis Programme on behalf of the National Institute for Health and Clinical Excellence – protocol: Title of project: Artificial Intelligence software for analysing chest X-ray images to identify suspected lung cancer*. 2022. URL: https://www.nice.org.uk/guidance/gid-dg10065/documents/final-protocol (Accessed 26 January 2023).
15. Tappenden P. *Conceptual modelling for health economic model development [HEDS Discussion Paper 12/05]*. White Rose Research Online; 2012. URL: https://eprints.whiterose.ac.uk/74464/1/HEDSDP1205.pdf (Accessed 28 November 2022).
16. Callister ME, Baldwin DR, Akram AR, Barnard S, Cane P, Draffan J*, et al.* British Thoracic Society guidelines for the investigation and management of pulmonary nodules. *Thorax* 2015;**70 Suppl 2**:ii1-ii54. http://dx.doi.org/10.1136/thoraxjnl-2015-207168

17. NHS England. *Updated National Optimal Lung Cancer Pathway – Sept 2020*. British Thoracic Oncology Group; 2020. URL: https://www.btog.org/news/just-released-updated-national-optimal-lung-cancer-pathway-sept-2020/ (Accessed 23 November 2022).
18. National Institute for Health and Care Excellence. *NICE health technology evaluations: the manual: Process and methods [PMG36]*. 2022. URL: https://www.nice.org.uk/process/pmg36/ (Accessed 23 November 2022).
19. York Health Economics Consortium. *Evidence Standards Framework for Digital Health Technologies: Cost Consequences and Budget Impact Analyses and Data Sources*. National Institute for Health and Care Excellence; 2021. URL: https://www.nice.org.uk/Media/Default/About/what-we-do/our-programmes/evidence-standards-framework/budget-impact-guide.pdf (Accessed 28 November 2022).
20. Mauskopf JA, Sullivan SD, Annemans L, Caro J, Mullins CD, Nuijten M*, et al.* Principles of good practice for budget impact analysis: report of the ISPOR Task Force on good research practices--budget impact analysis. *Value Health* 2007;**10**(5):336-47. http://dx.doi.org/10.1111/j.1524-4733.2007.00187.x
21. van Beek EJR, Ahn JS, Kim MJ, Murchison JT. Validation study of machine-learning chest radiograph software in primary and emergency medicine. *Clin Radiol* 2022;**25**:25. http://dx.doi.org/10.1016/j.crad.2022.08.129
22. Dissez G, Tay N, Dyer T, Tam M, Dittrich R, Doyne D*, et al. Enhancing Early Lung Cancer Detection on Chest Radiographs with AI-assistance: A Multi-Reader Study [preprint]*. arXiv.org; 2022. URL: https://arxiv.org/ftp/arxiv/papers/2208/2208.14742.pdf (Accessed 3 January 2023).
23. Nam JG, Hwang EJ, Kim DS, Yoo SJ, Choi H, Goo JM*, et al.* Undetected Lung Cancer at Posteroanterior Chest Radiography: Potential Role of a Deep Learning-based Detection Algorithm. *Radiol Cardiothorac Imaging* 2020;**2**(6):e190222. http://dx.doi.org/10.1148/ryct.2020190222
24. Jang S, Song H, Shin YJ, Kim J, Lee KW, Lee SS*, et al.* Deep Learning-based Automatic Detection Algorithm for Reducing Overlooked Lung Cancers on Chest Radiographs. *Radiology* 2020;**296(3)**:652-61. http://dx.doi.org/10.1148/radiol.2020200165
25. Koo YH, Shin KE, Park JS, Lee JW, Byun S, Lee H. Extravalidation and reproducibility results of a commercial deep learning-based automatic detection algorithm for pulmonary nodules on chest radiographs at tertiary hospital. *J Med Imaging Radiat Oncol* 2021;**65**(1):15-22. http://dx.doi.org/10.1111/1754-9485.13105
26. Homayounieh F, Digumarthy S, Ebrahimian S, Rueckel J, Hoppe BF, Sabel BO*, et al.* An Artificial Intelligence-Based Chest X-ray Model on Human Nodule Detection Accuracy From a Multicenter Study. *JAMA Network Open* 2021;**4**(12):e2141096. http://dx.doi.org/10.1001/jamanetworkopen.2021.41096
27. Gur D, Bandos AI, Cohen CS, Hakim CM, Hardesty LA, Ganott MA*, et al.* The "laboratory" effect: comparing radiologists' performance and variability during prospective clinical and laboratory mammography interpretations. *Radiology* 2008;**249**(1):47-53. http://dx.doi.org/10.1148/radiol.2491072025
28. Ferlay J, Ervik M, Lam F, Colombet M, Mery L, Piñeros M*, et al. Global Cancer Observatory: Cancer Today*. Lyon, France: International Agency for Research on Cancer; 2020. URL: https://gco.iarc.fr/today/home (Accessed 14 February 2023).
29. Lundh A, Lexchin J, Mintzes B, Schroll JB, Bero L. Industry sponsorship and research outcome. *Cochrane Database Syst Rev* 2017;**2**(2):Mr000033. http://dx.doi.org/10.1002/14651858.MR000033.pub3
30. Bajre MK, Pennington M, Woznitza N, Beardmore C, Radhakrishnan M, Harris R*, et al.* Expanding the role of radiographers in reporting suspected lung cancer: A cost-effectiveness analysis using a decision tree model. *Radiography (Lond)* 2017;**23**(4):273-8. http://dx.doi.org/10.1016/j.radi.2017.07.011

31. Snowsill T, Yang H, Griffin E, Long L, Varley-Campbell J, Coelho H*, et al.* Low-dose computed tomography for lung cancer screening in high-risk populations: a systematic review and economic evaluation. *Health Technol Assess* 2018;**22**(69):1-276. http://dx.doi.org/10.3310/hta22690
32. Exeter Test Group and Health Economics Group. *Interim report on the cost-effectiveness of low dose computed tomography (LDCT) screening for lung cancer in high risk individuals*. UK NSC Recommendations, GOV.UK; 2022. URL: https://view-health-screening-recommendations.service.gov.uk/document/586/download (Accessed 17 January 2023).
33. Foley RW, Nassour V, Oliver HC, Hall T, Masani V, Robinson G*, et al.* Chest X-ray in suspected lung cancer is harmful. *Eur Radiol* 2021;**31**(8):6269-74. http://dx.doi.org/10.1007/s00330-021-07708-0
34. Bradley SH, Bhartia BS, Callister ME, Hamilton WT, Hatton NLF, Kennedy MP*, et al.* Chest X-ray sensitivity and lung cancer outcomes: a retrospective observational study. *Br J Gen Pract* 2021;**71**(712):e862-e8. http://dx.doi.org/10.3399/bjgp.2020.1099
35. Woznitza N, Piper K, Rowe S, Bhowmik A. Immediate reporting of chest X-rays referred from general practice by reporting radiographers: a single centre feasibility study. *Clin Radiol* 2018;**73**(5):507.e1-.e8. http://dx.doi.org/10.1016/j.crad.2017.11.016
36. Woznitza N, Ghimire B, Devaraj A, Janes SM, Piper K, Rowe S*, et al.* Impact of radiographer immediate reporting of X-rays of the chest from general practice on the lung cancer pathway (radioX): a randomised controlled trial. *Thorax* 2022. http://dx.doi.org/10.1136/thorax-2022-219210
37. Dwyer-Hemmings L, Fairhead C. The diagnostic performance of chest radiographs for lung malignancy in symptomatic primary-care populations: A systematic review and meta-analysis. *BJR Open* 2021;**3**(1):20210005. http://dx.doi.org/10.1259/bjro.20210005
38. Naik H, Howell D, Su J, Qiu X, Brown C, Vennettilli A*, et al.* Stage specific health utility index scores of Canadian cancer patients. *J Clin Oncol* 2015;**33**(15 suppl):6614. http://dx.doi.org/10.1200/JCO.2015.33.15_SUPPL.6614
39. Bradley SH, Abraham S, Callister ME, Grice A, Hamilton WT, Lopez RR*, et al.* Sensitivity of chest X-ray for detecting lung cancer in people presenting with symptoms: a systematic review. *Br J Gen Pract* 2019;**69**(689):e827-e35. http://dx.doi.org/10.3399/bjgp19X706853
40. Lawson M, Underhill S, Chauhan M, Robinson S, Melesi V, Miller S*, et al.* P13 Mapping the lung cancer pathway. *Thorax* 2021;**76**(Suppl 1):A92-A. http://dx.doi.org/10.1136/thorax-2020-BTSabstracts.158
41. The Royal College of Radiologists. *Radiology reporting figures for service planning*. 2022. URL: https://www.rcr.ac.uk/system/files/publication/field_publication_files/radiology-reporting-figures-2022.pdf (Accessed 24 January 2023).
42. Evison M, Hewitt K, Lyons J, Crosbie P, Balata H, Gee C*, et al.* Implementation and outcomes of the RAPID programme: Addressing the front end of the lung cancer pathway in Manchester. *Clin Med (Northfield Il)* 2020;**20**(4):401-5. http://dx.doi.org/10.7861/clinmed.2019-0218
43. Byrne SC, Barrett B, Bhatia R. The impact of diagnostic imaging wait times on the prognosis of lung cancer. *Can Assoc Radiol J* 2015;**66**(1):53-7. http://dx.doi.org/10.1016/j.carj.2014.01.003
44. Peters JL, Snowsill TM, Griffin E, Robinson S, Hyde CJ. Variation in Model-Based Economic Evaluations of Low-Dose Computed Tomography Screening for Lung Cancer: A Methodological Review. *Value Health* 2022;**25**(4):656-65. http://dx.doi.org/10.1016/j.jval.2021.11.1352
45. Geppert J, Auguste P, Asgharzadeh A, Ghiasvand H, Patel M, Brown A*, et al. Software with artificial intelligence derived algorithms for automated detection and analysis of lung nodules from CT scan images: A Diagnostics Assessment Report*. National Institute for Health and Care Excellence; 2022. URL: https://www.nice.org.uk/guidance/GID-DG10046/documents/diagnostics-assessment-report (Accessed 14 February 2023).
46. Field JK, Duffy SW, Baldwin DR, Brain KE, Devaraj A, Eisen T*, et al.* The UK Lung Cancer Screening Trial: a pilot randomised controlled trial of low-dose computed tomography screening for the early detection of lung cancer. *Health Technol Assess* 2016;**20**(40):1-146. http://dx.doi.org/10.3310/hta20400

47. Horeweg N, van Rosmalen J, Heuvelmans MA, van der Aalst CM, Vliegenthart R, Scholten ET*, et al.* Lung cancer probability in patients with CT-detected pulmonary nodules: a prespecified analysis of data from the NELSON trial of low-dose CT screening. *Lancet Oncol* 2014;**15**(12):1332-41. http://dx.doi.org/10.1016/s1470-2045(14)70389-4
48. *National schedule of NHS costs: The main schedule, showing data for the whole range of services provided by provider, including admitted patient care on a finished consultant episode (FCE) basis.* NHS England; 2022. URL: https://www.england.nhs.uk/wp-content/uploads/2022/07/2_National_schedule_of_NHS_costs_FY20-21.xlsx (Accessed 14 February 2023).
49. Mazzone PJ, Obuchowski N, Fu AZ, Phillips M, Meziane M. Quality of life and healthcare use in a randomized controlled lung cancer screening study. *Ann Am Thorac Soc* 2013;**10**(4):324-9. http://dx.doi.org/10.1513/AnnalsATS.201301-007OC
50. Sutton AJ, Sagoo GS, Jackson L, Fisher M, Hamilton-Fairley G, Murray A*, et al.* Cost-effectiveness of a new autoantibody test added to Computed Tomography (CT) compared to CT surveillance alone in the diagnosis of lung cancer amongst patients with indeterminate pulmonary nodules. *PLoS One* 2020;**15**(9):e0237492. http://dx.doi.org/10.1371/journal.pone.0237492
51. Stevenson M, Lloyd-Jones M, Morgan MY, Wong R. Non-invasive diagnostic assessment tools for the detection of liver fibrosis in patients with suspected alcohol-related liver disease: a systematic review and economic evaluation. *Health Technol Assess* 2012;**16**(4):1-174. http://dx.doi.org/10.3310/hta16040
52. Jones K, Burns A. The NHS cost inflation index (NHSCII). In: *Unit costs of health and social care* Canterbury: Personal Social Services Research Unit, University of Kent; 2021. https://www.pssru.ac.uk/pub/uc/uc2021/sourcesofinformation.pdf
53. Lockwood P. An economic evaluation of introducing a skills mix approach to CT head reporting in clinical practice. *Radiography* 2016;**22**(2):124-30. http://dx.doi.org/10.1016/j.radi.2015.09.004
54. Jones K, Burns A. *Unit costs of health and social care*. Canterbury: Personal Social Services Research Unit, University of Kent; 2021. http://dx.doi.org/10.22024/UniKent/01.02.92342
55. Birtwistle M, Earnshaw A. *Saving lives, averting costs: an analysis of the financial implications of achieving earlier diagnosis of colorectal, lung and ovarian cancer*: Incisive Health, Cancer Research UK; 2014.
56. Kennedy MP, Hall PS, Callister ME. Factors affecting hospital costs in lung cancer patients in the United Kingdom. *Lung Cancer* 2016;**97**:8-14. http://dx.doi.org/10.1016/j.lungcan.2016.04.009
57. McGuire A, Martin M, Lenz C, Sollano JA. Treatment cost of non-small cell lung cancer in three European countries: comparisons across France, Germany, and England using administrative databases. *J Med Econ* 2015;**18**(7):525-32. http://dx.doi.org/10.3111/13696998.2015.1032974
58. Taylor-Phillips S, Geppert J, Stinton C, Freeman K, Johnson S, Fraser H*, et al.* Comparison of a full systematic review versus rapid review approaches to assess a newborn screening test for tyrosinemia type 1. *Res Synth Methods* 2017;**8**(4):475-84. http://dx.doi.org/10.1002/jrsm.1255
59. Hailey DM. Health technology assessment in Canada: diversity and evolution. *Med J Aust* 2007;**187**(5):286-8. http://dx.doi.org/10.5694/j.1326-5377.2007.tb01245.x
60. Watt A, Cameron A, Sturm L, Lathlean T, Babidge W, Blamey S*, et al.* Rapid reviews versus full systematic reviews: An inventory of current methods and practice in health technology assessment: Corrigendum. *Int J Technol Assess Health Care* 2008;**24**(03):369. http://dx.doi.org/10.1017/s0266462308080483

Appendices

Appendix 1: Literature searches

Test accuracy, practical implications, and clinical effectiveness

Table A1. Search strategy summary: Multi-stranded, targeted approach

Search #	Search	Sources
1	Intervention (AI and chest x-ray) AND Study type ('Reviews (best balance of sensitivity and specificity)' Clinical Queries limit OR systematic reviews filter (specific filter))	Epistemonikos, MEDLINE, Embase, CDSR, a computer science database
2	Intervention [broader] (AI) AND lung cancer or lung nodule AND study type (systematic reviews filter (specific filter))	Epistemonikos, MEDLINE, Embase, CDSR, a computer science database
3	Intervention (AI and chest x-ray) AND selected outcomes (lung cancer / lung nodule)	MEDLINE, Embase, CENTRAL (inc. trial register records), a computer science database
4	Technology names / companies [look in title, abstract and institution fields] AND (chest x-ray / lung cancer / lung nodule)	MEDLINE, Embase, CENTRAL (inc. trial register records), a computer science database
	Targeted searches for relevant ongoing systematic reviews	PROSPERO
	Targeted searches for relevant ongoing trials	WHO ICTRP
	Check references of relevant reviews and studies found via NICE and team members' scoping or clinical experts	NICE, EAG team members, clinical experts

Bibliographic databases

Table A2. Bibliographic databases: search summary

Source(s)	Date searched	Purpose	Description of search	Hits	Notes
MEDLINE (Ovid)	25/11/2022	Search to identify relevant reviews and primary studies	The 4 targeted searches run together (see table A1)	1119	Limited to English Language or no language specified. Non-human studies, letters, editorials, comments removed. No date limits applied.
Embase (Ovid)	29/11/2022	Search to identify relevant reviews and primary studies	The 4 targeted searches run together (see table A1)	2198	Limited to English Language or no language specified. Non-human studies, letters, editorials, removed. No date limits applied.
Cochrane Database of Systematic Reviews (Wiley)	30/11/2022	Search to identify relevant reviews for reference checking	Intervention (AI and chest x-ray) OR Intervention [broader] (AI) AND lung cancer / lung nodule	0	Specialist database for Cochrane systematic reviews
Cochrane CENTRAL (Wiley)	30/11/2022	Search to identify relevant primary studies	Intervention (AI and chest x-ray) AND lung cancer or lung nodule OR Technology names / companies AND (chest x-ray / lung cancer / lung nodule)	52	Specialist database for trials No date or language limits applied.
Epistemonikos	01/12/2022	Search to identify relevant reviews for reference checking	Intervention (AI and chest x-ray) OR Intervention [broader] (AI) AND	45	Specialist database for systematic reviews and overviews.

			lung cancer / lung nodule		Filtered for publication types: • Systematic Review • Broad Synthesis No date or language limits applied.
ACM Digital Library	01/12/2022	Search to identify relevant reviews and primary studies in a computer science database	Intervention (AI and chest x-ray) OR Intervention [broader] (AI) AND lung cancer / lung nodule)	12	Limited to Content Type: Review article No date or language limits applied.
			Intervention (AI and chest x-ray) AND lung cancer / lung nodule	452	No limits applied.
			Technology names / companies AND (chest x-ray / lung cancer / lung nodule	1	No limits applied.

Totals

Total from databases: 3879
Total after duplicates removed: 3049

Medline (Ovid)
Searched 25/11/2022
Ovid MEDLINE® ALL <1946 to November 23, 2022>

1 exp artificial intelligence/ or exp machine learning/ or exp deep learning/ or exp supervised machine learning/ or exp support vector machine/ or exp unsupervised machine learning/ 160931
2 ai.kf,tw.39919
3 ((artificial or machine or deep) adj5 (intelligence or learning or reasoning)).kf,tw.124190
4 exp Neural Networks, Computer/ 53917
5 (neural network* or convolutional or CNN or CNNs).kf,tw. 90349
6 exp Diagnosis, Computer-Assisted/ 86384
7 Pattern Recognition, Automated/ 26362
8 ((automat* or autonomous or computer aided or computer assisted) adj3 (detect* or identif* or diagnos*)).kf,tw. 33565
9 (support vector machine* or random forest* or black box learning).kf,tw. 37636
10 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 [AI] 396688

11 exp Radiography, Thoracic/ 40528
12 X-Rays/ 31129
13 (((chest or lung* or thora*) adj3 (radiograph* or radiogram* or radiology or roentgen* or x-ray* or xray* or film*)) or CXR*).kf,tw. 66459
14 11 or 12 or 13 [CXR] 121772
15 10 and 14 [AI and CXR] 3865
16 limit 15 to "reviews (best balance of sensitivity and specificity)" [AI and CXR and Reviews] 349
17 (metaanalys* or meta analys* or NMA* or MAIC* or indirect comparison* or mixed treatment comparison*).mp. 288007
18 (systematic* adj3 (review* or overview* or search or literature)).mp. 328557
19 17 or 18 459498
20 15 and 19 [AI and CXR and SRs] 40
21 16 or 20 [AI and CXR and Reviews / SRs] 360
22 exp Lung Neoplasms/ or Solitary Pulmonary Nodule/ 268336
23 ((lung or lungs or pulmon* or intrapulmon* or bronch*) adj3 (abnormal* or nodul* or lesion* or mass or masses or cancer* or neoplas* or tumor* or tumour* or carcino* or malignan* or adenocarcinom* or blastoma*)).kf,tw. 326364
24 ((pancoast* or superior sulcus or pulmonary sulcus) adj4 (tumor* or tumour* or syndrome*)).kf,tw. 946
25 (sclc or nsclc).kf,tw. 64440
26 22 or 23 or 24 or 25 [Lung Cancer / Nodule] 398150
27 10 and 26 [AI and Lung Cancer / Nodule] 6749
28 (metaanalys* or meta analys* or NMA* or MAIC* or indirect comparison* or mixed treatment comparison*).mp. 288007
29 (systematic* adj3 (review* or overview* or search or literature)).mp. 328557
30 28 or 29 [SRs] 459498
31 27 and 30 [AI and Lung Cancer / Nodule and SRs] 100
32 10 and 14 and 26 [AI and CXR and Lung Cancer / Nodule] 707
33 AI-Rad Companion Chest X-ray*.kf,tw,in. 1
34 Annalise CXR*.kf,tw,in. 1
35 Auto Lung Nodule Detection*.kf,tw,in. 0
36 ChestView*.kf,tw,in. 0
37 (Chest X-Ray Classifier* or Quibim*).kf,tw,in. 46
38 CheXVision*.kf,tw,in. 0
39 (ClearRead Xray* adj2 Detect).kf,tw,in. 0
40 InferRead DR Chest*.kf,tw,in. 0
41 JLD-02K*.kf,tw,in. 0
42 Lunit INSIGHT CXR*.kf,tw,in. 4
43 Milvue Suite*.kf,tw,in. 0
44 ChestEye Quality*.kf,tw,in. 0
45 (qXR* or Qure*).kf,tw,in. 6815
46 (red dot* or behold*).kf,tw,in. 1090
47 SenseCare-Chest DR Pro*.kf,tw,in. 0
48 VUNO Med-Chest X-Ray*.kf,tw,in. 0
49 (X1* and Visionairy Health).kf,tw,in. 0
50 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 [Technology Names / Companies] 7956
51 50 and 14 [Technology Names / Companies and CXR] 61
52 50 and 26 [Technology Names / Companies and Lung Cancer / Nodules] 90
53 51 or 52 [Technology Names / Companies and CXR / Lung Cancer / Nodules] 136

54 21 or 31 or 32 or 53 1190
55 limit 54 to english language 1134
56 limit 54 to no language specified 0
57 55 or 56 1134
58 exp animals/ not humans.sh. 5066999
59 57 not 58 1128
60 limit 59 to (comment or editorial or letter) 9
61 59 not 60 1119

Embase (Ovid)
Searched 29/11/2022
Embase Classic+Embase <1947 to 2022 Week 47>

1 exp artificial intelligence/ or exp machine learning/ 373033
2 ai.kf,tw.55274
3 ((artificial or machine or deep) adj5 (intelligence or learning or reasoning)).kf,tw.146615
4 (neural network* or convolutional or CNN or CNNs).kf,tw. 108457
5 computer assisted diagnosis/ or computer assisted radiography/44996
6 ((automat* or autonomous or computer aided or computer assisted) adj3 (detect* or identif* or diagnos*)).kf,tw. 44987
7 (support vector machine* or random forest* or black box learning).kf,tw. 46703
8 1 or 2 or 3 or 4 or 5 or 6 or 7 [AI] 530438
9 exp thorax radiography/ 230425
10 X ray/ 119143
11 (((chest or lung* or thora*) adj3 (radiograph* or radiogram* or radiology or roentgen* or x-ray* or xray* or film*)) or CXR*).kf,tw. 107803
12 9 or 10 or 11 [CXR] 379945
13 8 and 12 [AI and CXR] 5577
14 limit 13 to "reviews (best balance of sensitivity and specificity)" [AI and CXR and Reviews] 657
15 (metaanalys* or meta analys* or NMA* or MAIC* or indirect comparison* or mixed treatment comparison*).mp. 414514
16 (systematic* adj3 (review* or overview* or search or literature)).mp. 520359
17 15 or 16 695121
18 13 and 17 [AI and CXR and SRs] 117
19 14 or 18 [AI and CXR and Reviews / SRs] 678
20 exp lung tumor/ or lung nodule/ 495858
21 ((lung or lungs or pulmon* or intrapulmon* or bronch*) adj3 (abnormal* or nodul* or lesion* or mass or masses or cancer* or neoplas* or tumor* or tumour* or carcino* or malignan* or adenocarcinom* or blastoma*)).kf,tw. 493166
22 ((pancoast* or superior sulcus or pulmonary sulcus) adj4 (tumor* or tumour* or syndrome*)).kf,tw. 1328
23 (sclc or nsclc).kf,tw. 116762
24 20 or 21 or 22 or 23 [Lung Cancer / Nodule] 655493
25 8 and 24 [AI and Lung Cancer / Nodule] 12931
26 (metaanalys* or meta analys* or NMA* or MAIC* or indirect comparison* or mixed treatment comparison*).mp. 414514
27 (systematic* adj3 (review* or overview* or search or literature)).mp. 520359
28 26 or 27 [SRs] 695121
29 25 and 28 [AI and Lung Cancer / Nodule and SRs] 313
30 8 and 12 and 24 [AI and CXR and Lung Cancer / Nodule] 1114

31 AI-Rad Companion Chest X-ray*.kf,tw,in. 1
32 Annalise CXR*.kf,tw,in. 1
33 Auto Lung Nodule Detection*.kf,tw,in. 0
34 ChestView*.kf,tw,in. 0
35 (Chest X-Ray Classifier* or Quibim*).kf,tw,in. 57
36 CheXVision*.kf,tw,in. 0
37 (ClearRead Xray* adj2 Detect).kf,tw,in. 0
38 InferRead DR Chest*.kf,tw,in. 0
39 JLD-02K*.kf,tw,in. 0
40 Lunit INSIGHT CXR*.kf,tw,in. 6
41 Milvue Suite*.kf,tw,in. 0
42 ChestEye Quality*.kf,tw,in. 0
43 (qXR* or Qure*).kf,tw,in. 14268
44 (red dot* or behold*).kf,tw,in. 1520
45 SenseCare-Chest DR Pro*.kf,tw,in. 0
46 VUNO Med-Chest X-Ray*.kf,tw,in. 0
47 (X1* and Visionairy Health).kf,tw,in. 0
48 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 [Technology Names / Companies] 15850
49 48 and 12 [Technology Names / Companies and CXR] 267
50 48 and 24 [Technology Names / Companies and Lung Cancer / Nodules] 234
51 49 or 50 [Technology Names / Companies and CXR / Lung Cancer / Nodules] 466
52 19 or 29 or 30 or 51 2362
53 limit 52 to english language 2271
54 limit 52 to no language specified 1
55 53 or 54 2272
56 animal experiment/ not (human experiment/ or human/) 2472698
57 55 not 56 2263
58 limit 57 to (editorial or letter) 65
59 57 not 58 2198

Cochrane Database of Systematic Reviews (Wiley)
Search Name: qXR EVA Reviews
Date Run: 30/11/2022 19:30:29
Comment: 30Nov2022

ID Search Hits
#1 [mh "artificial intelligence"] OR [mh "machine learning"] OR [mh "deep learning"] OR [mh "supervised machine learning"] OR [mh "support vector machine"] OR [mh "unsupervised machine learning"] 1540
#2 ai:ti,ab,kw 5002
#3 ((artificial OR machine OR deep) NEAR/5 (intelligence OR learning OR reasoning)):ti,ab,kw 3847
#4 [mh "Neural Networks, Computer"] 217
#5 (("neural" NEXT network*) OR convolutional OR CNN OR CNNs):ti,ab,kw 1738
#6 [mh "Diagnosis, Computer-Assisted"] 1943
#7 [mh ^"Pattern Recognition, Automated"] 193
#8 ((automat* OR autonomous OR "computer aided" OR "computer assisted") NEAR/3 (detect* OR identif* OR diagnos*)):ti,ab,kw 2092

#9 (("support vector" NEXT machine*) OR ("random" NEXT forest*) OR "black box learning"):ti,ab,kw 935
#10 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 13357
#11 [mh "Radiography, Thoracic"] 363
#12 [mh ^X-Rays] 59
#13 (((chest OR lung* OR thora*) NEAR/3 (radiograph* OR radiogram* OR radiology OR roentgen* OR x-ray* OR xray* OR film*)) OR CXR*):ti,ab,kw 5878
#14 #11 OR #12 OR #13 5948
#15 #10 AND #14 120
#16 [mh "Lung Neoplasms"] OR [mh ^"Solitary Pulmonary Nodule"] 8755
#17 ((lung OR lungs OR pulmon* OR intrapulmon* OR bronch*) NEAR/3 (abnormal* OR nodul* OR lesion* OR mass OR masses OR cancer* OR neoplas* OR tumor* OR tumour* OR carcino* OR malignan* OR adenocarcinom* OR blastoma*)):ti,ab,kw 28597
#18 ((pancoast* OR "superior sulcus" OR "pulmonary sulcus") NEAR/4 (tumor* OR tumour* OR syndrome*)):ti,ab,kw 17
#19 (sclc OR nsclc):ti,ab,kw 12248
#20 #16 OR #17 OR #18 OR #19 29193
#21 #10 AND #20 348
#22 #15 OR #21 421

Cochrane Reviews: 0

CENTRAL (Wiley)

Search Name: qXR EVA Trials
Date Run: 30/11/2022 22:52:13
Comment: 30Nov2022

ID Search Hits
#1 [mh "artificial intelligence"] OR [mh "machine learning"] OR [mh "deep learning"] OR [mh "supervised machine learning"] OR [mh "support vector machine"] OR [mh "unsupervised machine learning"] 1540
#2 ai:ti,ab,kw 5002
#3 ((artificial OR machine OR deep) NEAR/5 (intelligence OR learning OR reasoning)):ti,ab,kw 3847
#4 [mh "Neural Networks, Computer"] 217
#5 (("neural" NEXT network*) OR convolutional OR CNN OR CNNs):ti,ab,kw 1738
#6 [mh "Diagnosis, Computer-Assisted"] 1943
#7 [mh ^"Pattern Recognition, Automated"] 193
#8 ((automat* OR autonomous OR "computer aided" OR "computer assisted") NEAR/3 (detect* OR identif*OR diagnos*)):ti,ab,kw 2092
#9 (("support vector" NEXT machine*) OR ("random" NEXT forest*) OR "black box learning"):ti,ab,kw 935
#10 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 13357
#11 [mh "Radiography, Thoracic"] 363
#12 [mh ^X-Rays] 59
#13 ((chest OR lung* OR thora*) NEAR/3 (radiograph* OR radiogram* OR radiology OR roentgen* OR x-ray* OR xray* OR film* OR CXR*)):ti,ab,kw 5878
#14 #11 OR #12 OR #13 5948
#15 [mh "Lung Neoplasms"] OR [mh ^"Solitary Pulmonary Nodule"] 8755

#16 ((lung OR lungs OR pulmon* OR intrapulmon* OR bronch*) NEAR/3 (abnormal* OR nodul* OR lesion* OR mass OR masses OR cancer* OR neoplas* OR tumor* OR tumour* OR carcino* OR malignan* OR adenocarcinom* OR blastoma*)):ti,ab,kw 28597
#17 ((pancoast* OR "superior sulcus" OR "pulmonary sulcus") NEAR/4 (tumor* OR tumour* OR syndrome*)):ti,ab,kw 17
#18 (sclc OR nsclc):ti,ab,kw 12248
#19 #15 OR #16 OR #17 OR #18 29193
#20 #10 and #14 and #19 47
#21 ("AI-Rad Companion Chest" NEXT X-ray*) 0
#22 ("Annalise" NEXT CXR*) 0
#23 ("Auto Lung Nodule" NEXT Detection*) 0
#24 ChestView* 0
#25 (("Chest X-Ray" NEXT Classifier*) OR Quibim*) 0
#26 CheXVision* 0
#27 (("ClearRead" NEXT Xray*) NEAR/2 Detect) 0
#28 ("InferRead DR" NEXT Chest*) 0
#29 JLD-02K* 0
#30 ("Lunit INSIGHT" NEXT CXR*) 2
#31 ("Milvue" NEXT Suite*) 0
#32 ("ChestEye" NEXT Quality*) 0
#33 (qXR* OR Qure*) 921
#34 (("red" NEXT dot*) OR behold*) 71
#35 ("SenseCare-Chest DR" NEXT Pro*) 0
#36 ("VUNO Med-Chest" NEXT X-Ray*) 1
#37 (X1* AND "Visionairy Health") 0
#38 #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37 995
#39 #14 and #38 4
#40 #19 and #38 7
#41 #39 or #40 8
#42 #20 or #41 53

Trials: 52

Epistemonikos
Searched 01/12/2022

(title:(("AI" OR "artificial intelligence" OR "artificial learning" OR "artificial reasoning" OR "machine intelligence" OR "machine learning" OR "machine reasoning" OR "deep intelligence" OR "deep learning" OR "deep reasoning" OR "neural network" OR "neural networks" OR "neural networking" OR convolutional OR "CNN" OR "CNNs" OR ((automat* OR autonomous OR "computer aided" OR "computer assisted") AND (detect* OR identif* OR diagnos*)) OR "support vector machine" OR "support vector machines" OR "support vector network" OR "support vector networks" OR "random forest" OR "random forests" OR "black box learning") AND ((((chest OR lung* OR thora*) AND (radiograph* OR radiogram* OR radiology OR roentgen* OR x-ray* OR xray* OR film*)) OR CXR*) OR ((lung OR lungs OR pulmon* OR intrapulmon* OR bronch*) AND (abnormal* OR nodul* OR lesion* OR mass OR masses OR cancer* OR neoplas* OR tumor* OR tumour* OR carcino* OR malignan* OR adenocarcinom* OR blastoma*)) OR ((pancoast* OR superior sulcus OR pulmonary sulcus) AND (tumor* OR tumour* OR syndrome*)))) OR abstract:(("AI" OR "artificial intelligence" OR "artificial learning" OR "artificial reasoning" OR "machine intelligence" OR "machine learning" OR "machine

reasoning" OR "deep intelligence" OR "deep learning" OR "deep reasoning" OR "neural network" OR "neural networks" OR "neural networking" OR convolutional OR "CNN" OR "CNNs" OR ((automat* OR autonomous OR "computer aided" OR "computer assisted") AND (detect* OR identif* OR diagnos*)) OR "support vector machine" OR "support vector machines" OR "support vector network" OR "support vector networks" OR "random forest" OR "random forests" OR "black box learning") AND ((((chest OR lung* OR thora*) AND (radiograph* OR radiogram* OR radiology OR roentgen* OR x-ray* OR xray* OR film*)) OR CXR*) OR ((lung OR lungs OR pulmon* OR intrapulmon* OR bronch*) AND (abnormal* OR nodul* OR lesion* OR mass OR masses OR cancer* OR neoplas* OR tumor* OR tumour* OR carcino* OR malignan* OR adenocarcinom* OR blastoma*)) OR ((pancoast* OR superior sulcus OR pulmonary sulcus) AND (tumor* OR tumour* OR syndrome*)))))

Publication type:
Systematic Review: 44
Broad Synthesis: 1

ACM Digital Library
Searched 01/12/2022

Search for reviews

https://dl.acm.org/search/advanced

Selected ACM Guide to Computing Literature

Title:((("AI" OR "artificial intelligence" OR "artificial learning" OR "artificial reasoning" OR "machine intelligence" OR "machine learning" OR "machine reasoning" OR "deep intelligence" OR "deep learning" OR "deep reasoning" OR "neural network" OR "neural networks" OR "neural networking" OR convolutional OR "CNN" OR "CNNs" OR (automat* OR autonomous OR "computer aided" OR "computer assisted") AND (detect* OR identif* OR diagnos*) OR "support vector machine" OR "support vector machines" OR "support vector network" OR "support vector networks" OR "random forest" OR "random forests" OR "black box learning") AND ((((chest OR lung* OR thora*) AND (radiograph* OR radiogram* OR radiology OR roentgen* OR x-ray* OR xray* OR film*)) OR CXR*) OR ((lung OR lungs OR pulmon* OR intrapulmon* OR bronch*) AND (abnormal* OR nodul* OR lesion* OR mass OR masses OR cancer* OR neoplas* OR tumor* OR tumour* OR carcino* OR malignan* OR adenocarcinom* OR blastoma*)) OR ((pancoast* OR "superior sulcus" OR "pulmonary sulcus") AND (tumor* OR tumour* OR syndrome*))))) OR Abstract:((("AI" OR "artificial intelligence" OR "artificial learning" OR "artificial reasoning" OR "machine intelligence" OR "machine learning" OR "machine reasoning" OR "deep intelligence" OR "deep learning" OR "deep reasoning" OR "neural network" OR "neural networks" OR "neural networking" OR convolutional OR "CNN" OR "CNNs" OR (automat* OR autonomous OR "computer aided" OR "computer assisted") AND (detect* OR identif* OR diagnos*) OR "support vector machine" OR "support vector machines" OR "support vector network" OR "support vector networks" OR "random forest" OR "random forests" OR "black box learning") AND ((((chest OR lung* OR thora*) AND (radiograph* OR radiogram* OR radiology OR roentgen* OR x-ray* OR xray* OR film*)) OR CXR*) OR ((lung OR lungs OR pulmon* OR intrapulmon* OR bronch*) AND (abnormal* OR nodul* OR lesion* OR mass OR masses OR cancer* OR neoplas* OR tumor* OR tumour* OR carcino* OR malignan* OR adenocarcinom* OR blastoma*)) OR ((pancoast* OR "superior sulcus" OR "pulmonary sulcus") AND (tumor* OR tumour* OR syndrome*)))))

Filter by
Content type:

Review Article: 12

Searches for primary studies

Searched 01/12/2022
https://dl.acm.org/search/advanced
Selected ACM Guide to Computing Literature

Title:((("AI" OR "artificial intelligence" OR "artificial learning" OR "artificial reasoning" OR "machine intelligence" OR "machine learning" OR "machine reasoning" OR "deep intelligence" OR "deep learning" OR "deep reasoning" OR "neural network" OR "neural networks" OR "neural networking" OR convolutional OR "CNN" OR "CNNs" OR (automat* OR autonomous OR "computer aided" OR "computer assisted") AND (detect* OR identif* OR diagnos*) OR "support vector machine" OR "support vector machines" OR "support vector network" OR "support vector networks" OR "random forest" OR "random forests" OR "black box learning") AND (((chest OR lung* OR thora*) AND (radiograph* OR radiogram* OR radiology OR roentgen* OR x-ray* OR xray* OR film*)) OR CXR*) AND ((lung OR lungs OR pulmon* OR intrapulmon* OR bronch*) AND (abnormal* OR nodul* OR lesion* OR mass OR masses OR cancer* OR neoplas* OR tumor* OR tumour* OR carcino* OR malignan* OR adenocarcinom* OR blastoma*)) OR ((pancoast* OR "superior sulcus" OR "pulmonary sulcus") AND (tumor* OR tumour* OR syndrome*)))) OR Abstract:((("AI" OR "artificial intelligence" OR "artificial learning" OR "artificial reasoning" OR "machine intelligence" OR "machine learning" OR "machine reasoning" OR "deep intelligence" OR "deep learning" OR "deep reasoning" OR "neural network" OR "neural networks" OR "neural networking" OR convolutional OR "CNN" OR "CNNs" OR (automat* OR autonomous OR "computer aided" OR "computer assisted") AND (detect* OR identif* OR diagnos*) OR "support vector machine" OR "support vector machines" OR "support vector network" OR "support vector networks" OR "random forest" OR "random forests" OR "black box learning") AND (((chest OR lung* OR thora*) AND (radiograph* OR radiogram* OR radiology OR roentgen* OR x-ray* OR xray* OR film*)) OR CXR*) AND ((lung OR lungs OR pulmon* OR intrapulmon* OR bronch*) AND (abnormal* OR nodul* OR lesion* OR mass OR masses OR cancer* OR neoplas* OR tumor* OR tumour* OR carcino* OR malignan* OR adenocarcinom* OR blastoma*)) OR ((pancoast* OR "superior sulcus" OR "pulmonary sulcus") AND (tumor* OR tumour* OR syndrome*))))

452

Searched 02/12/2022
https://dl.acm.org/search/advanced
Selected ACM Guide to Computing Literature

Title:(((ChestView* OR “Chest X-Ray Classifier” OR Quibim* OR CheXVision* OR (“ClearRead Xray” AND Detect) OR “InferRead DR Chest” OR JLD-02K* OR “Lunit INSIGHT CXR” OR “Milvue Suite” OR “ChestEye Quality” OR qXR* OR Qure* OR “red dot” or behold* OR “SenseCare-Chest DR Pro” OR “VUNO Med-Chest X-Ray” OR (X1* AND “Visionairy Health”)) AND ((((chest OR lung* OR thora*) AND (radiograph* OR radiogram* OR radiology OR roentgen* OR x-ray* OR xray* OR film*)) OR CXR*) OR ((lung OR lungs OR pulmon* OR intrapulmon* OR bronch*) AND (abnormal* OR nodul* OR lesion* OR mass OR masses OR cancer* OR neoplas* OR tumor* OR tumour* OR carcino* OR malignan* OR adenocarcinom* OR blastoma*)) OR ((pancoast* OR "superior sulcus" OR "pulmonary sulcus") AND (tumor* OR tumour* OR syndrome*))))) OR Abstract:(((ChestView* OR “Chest X-Ray Classifier” OR Quibim* OR CheXVision* OR (“ClearRead Xray” AND Detect) OR “InferRead DR Chest” OR JLD-02K*

OR “Lunit INSIGHT CXR” OR “Milvue Suite” OR “ChestEye Quality” OR qXR* OR Qure* OR “red dot” or behold* OR “SenseCare-Chest DR Pro” OR “VUNO Med-Chest X-Ray” OR (X1* AND “Visionairy Health”)) AND ((((chest OR lung* OR thora*) AND (radiograph* OR radiogram* OR radiology OR roentgen* OR x-ray* OR xray* OR film*)) OR CXR*) OR ((lung OR lungs OR pulmon* OR intrapulmon* OR bronch*) AND (abnormal* OR nodul* OR lesion* OR mass OR masses OR cancer* OR neoplas* OR tumor* OR tumour* OR carcino* OR malignan* OR adenocarcinom* OR blastoma*)) OR ((pancoast* OR "superior sulcus" OR "pulmonary sulcus") AND (tumor* OR tumour* OR syndrome*)))))

1

Systematic review register: search summary

PROSPERO
Searched 15/12/2022

#1 MeSH DESCRIPTOR Artificial Intelligence EXPLODE ALL TREES 477
#2 MeSH DESCRIPTOR machine learning EXPLODE ALL TREES 154
#3 MeSH DESCRIPTOR deep learning EXPLODE ALL TREES 23
#4 MeSH DESCRIPTOR supervised machine learning EXPLODE ALL TREES 1
#5 MeSH DESCRIPTOR support vector machine EXPLODE ALL TREES 0
#6 MeSH DESCRIPTOR unsupervised machine learning EXPLODE ALL TREES 0
#7 ai 1818
#8 (artificial or machine or deep) AND (intelligence or learning or reasoning) 1830
#9 MeSH DESCRIPTOR Neural Networks, Computer EXPLODE ALL TREES 28
#10 "neural network" or "neural networks" or convolutional or CNN or CNNs 481
#11 MeSH DESCRIPTOR Diagnosis, Computer-Assisted EXPLODE ALL TREES 15
#12 MeSH DESCRIPTOR Pattern Recognition, Automated EXPLODE ALL TREES 1
#13 ((automat* or autonomous or "computer aided" or "computer assisted") AND (detect* or identif* or diagnos*)) 3779
#14 "support vector machine" or "support vector machines" or "random forest" or "black box learning" 156
#15 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 6790
#16 MeSH DESCRIPTOR Radiography, Thoracic EXPLODE ALL TREES 10
#17 MeSH DESCRIPTOR X-Rays 29
#18 ((chest or lung* or thora*) and (radiograph* or radiogram* or radiology or roentgen* or x-ray* or xray* or film*)) or CXR* 1104
#19 #18 OR #17 OR #16 1120
#20 #15 AND #19 96
#21 MeSH DESCRIPTOR Lung Neoplasms EXPLODE ALL TREES 572
#22 MeSH DESCRIPTOR Solitary Pulmonary Nodule 6
#23 (lung or lungs or pulmon* or intrapulmon* or bronch*) AND (abnormal* or nodul* or lesion* or mass or masses or cancer* or neoplas* or tumor* or tumour* or carcino* or malignan* or adenocarcinom* or blastoma*) 6014
#24 (pancoast* or "superior sulcus" or "pulmonary sulcus") and (tumor* or tumour* or syndrome*) 5
#25 sclc or nsclc 896

#26 #21 OR #22 OR #23 OR #24 OR #25 6062
#27 #26 AND #15 256
#28 #27 OR #20 312
#29 #15 AND #19 AND #26 40

40 sifted online, 2 potentially relevant records sent to reviewers for checking

Trials registers: search summary

WHO ICTRP
Searched 18/01/2023 – targeted search #1

((lung* OR pulmonary OR intrapulmon* or bronch*) AND (abnormal* or nodul* or lesion* or mass or masses or cancer* or neoplas* or tumor* or tumour* or carcino* or malignan* or adenocarcinom* or blastoma*)) in the Condition

AND

(((artificial or machine or deep) AND (intelligence or learning or reasoning)) OR (AI OR "neural network*" or convolutional or CNN or CNNs OR "support vector machine*" or "random forest*" or "black box learning") OR ((automat* or autonomous or "computer aided" or "computer assisted") AND (detect* or identif* or diagnos*))) in the Intervention

AND

Recruitment status is All

32 records for 31 trials found

Searched 18/01/2023 – targeted search #2

((((artificial or machine or deep) AND (intelligence or learning or reasoning)) OR (AI OR "neural network*" or convolutional or CNN or CNNs OR "support vector machine*" or "random forest*" or "black box learning") OR ((automat* or autonomous or "computer aided" or "computer assisted") AND (detect* or identif* or diagnos*))) AND (((chest OR lung* OR thora*) AND (radiograph* OR radiogram* OR radiology OR roentgen* OR x-ray* OR xray* OR film*)) OR CXR*)) in the Intervention

13 records for 13 trials found

After de-duplicating with above:

12 records remaining

Searched 18/01/2023 – targeted search #3

((((artificial or machine or deep) AND (intelligence or learning or reasoning)) OR (AI OR "neural network*" or convolutional or CNN or CNNs OR "support vector machine*" or "random forest*" or "black box learning") OR ((automat* or autonomous or "computer aided" or "computer assisted") AND (detect* or identif* or diagnos*))) AND (((chest OR lung* OR thora*) AND (radiograph* OR radiogram* OR radiology OR roentgen* OR x-ray* OR xray* OR film*)) OR CXR*)) in the Title

29 records for 29 trials found

After de-duplicating with above: 22

Total from the 3 searches: 65
65 filtered by the information specialist for basic eligibility (CXR and lung cancer/nodule/abnormality, or unclear) or duplication with trial records found via other sources. 9 sent to clinical effectiveness reviewer for checking.

Cost-effectiveness searches

CEA Registry

Searched 30/11/2022

https://cear.tuftsmedicalcenter.org/

Basic Search

Keyword is: lung cancer

Total: 285

Basic Search

ICD-10: Malignant neoplasms of respiratory and intrathoracic organs (C30-C39)

Total: 264

Deduplicated in Excel

Copied and pasted results from second search into same sheet as the first search then…

using Home > Conditional Formatting > Highlight Cells Rules > Duplicate Values

… and scanned by eye for any unique references in the second search. Kept these and deleted the duplicates.

Total after deduplication: 303

MEDLINE (Ovid)

Searched 07/12/2022

Ovid MEDLINE(R) ALL <1946 to December 06, 2022>

1	exp Radiography, Thoracic/	40535
2	X-Rays/	31182
3	((chest or lung* or thora*) adj3 (radiograph* or radiogram* or radiology or roentgen* or x-ray* or xray*)).kf,tw.	64896
4	1 or 2 or 3	120457
5	exp Economics/	653642
6	exp "Costs and Cost Analysis"/	261580

7 Health Status/ 88924

8 exp "Quality of Life"/ 255297

9 exp Quality-Adjusted Life Years/15263

10 (pharmacoeconomic* or pharmaco-economic* or economic* or cost* or price or prices or pricing).ti,ab,kf. 1054159

11 (expenditure$ not energy).ti,ab,kf. 36095

12 (value adj1 money).ti,ab,kf. 40

13 budget*.ti,ab,kf. 34691

14 (health state* or health status).ti,ab,kf. 78185

15 (qaly* or ICER or utilit* or EQ5D or EQ-5D or euroqol or euro-qol or short-form 36 or shortform 36 or SF-36 or SF36 or SF-6D or SF6D or SF-12 or SF12 or health utilities index or HUI).ti,ab,kf. 311371

16 (markov or time trade off or TTO or standard gamble or SG or hrql or hrqol or disabilit* or disutilit* or net benefit or contingent valuation).ti,ab,kf. 302967

17 (quality adj2 life).ti,ab,kf. 364802

18 (decision adj2 model).ti,ab,kf. 8899

19 (visual analog* scale* or discrete choice experiment* or health* year* equivalen* or (willing* adj2 pay)).ti,ab,kf. 81000

20 resource*.ti,ab,kf. 447554

21 (well-being or wellbeing).ti,ab,kf. 130164

22 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 2829367

23 exp Lung Neoplasms/ or Solitary Pulmonary Nodule/ 268862

24 ((lung or lungs or pulmon* or intrapulmon* or bronch*) adj3 (abnormal* or nodul* or lesion* or mass or masses or cancer* or neoplas* or tumor* or tumour* or carcino* or malignan* or adenocarcinom* or blastoma*)).kf,tw. 327230

25 ((pancoast* or superior sulcus or pulmonary sulcus) adj4 (tumor* or tumour* or syndrome*)).kf,tw. 946

26 (sclc or nsclc).kf,tw. 64690

27 23 or 24 or 25 or 26 399076

28 4 and 22 and 27 817

Appendix 2: Table of studies excluded at full text assessment

Cancer studies

Reference	Reason for exclusion	Comments
Ahn JS, Ebrahimian S, McDermott S, Lee S, Naccarato L, Di Capua JF*, et al.* Association of Artificial Intelligence-Aided Chest Radiograph Interpretation with Reader Performance and Efficiency. *JAMA Network Open* 2022;**5(8)**:E2229289. http://dx.doi.org/10.1001/jamanetworkopen.2022.29289	Population	Software eligible. Radiologist + others using AI and without AI (4 weeks apart), CXR were from two hospital databases (one is an intensive care database) and no details of the referral route of participants or prior cancer status, included nodules as an outcome
Ajmera P, Pant R, Seth J, Ghuwalewala S, Kathuria S, Rathi S*, et al.* Deep-learning-based automatic detection of pulmonary nodules from chest radiographs. *medRxiv* 2022;**23**. http://dx.doi.org/10.1101/2022.06.21.22276691	Intervention	Not a named intervention, CXRs were from tertiary setting with no further details
Aoki T, Oda N, Yamashita Y, Yamamoto K, Korogi Y. Usefulness of computerized method for lung nodule detection on digital chest radiographs using similar subtraction images from different patients. *Eur J Radiol* 2012;**81**(5):1062-7. http://dx.doi.org/10.1016/j.ejrad.2011.02.010	Intervention	Software not stated, doesn't appear to be AI
Bae K, Oh DY, Yun ID, Jeon KN. Bone Suppression on Chest Radiographs for Pulmonary Nodule Detection: Comparison between a Generative Adversarial Network and Dual-Energy Subtraction. *Korean Journal of Radiology* 2022;**23**(1):139-49. http://dx.doi.org/10.3348/kjr.2021.0146	Intervention	Not a named intervention (and no AI component)
Baltruschat IM, Nickisch H, Grass M, Knopp T, Saalbach A. Comparison of Deep Learning Approaches for Multi-Label Chest X-Ray Classification. *Sci Rep* 2019;**9**(1):6381. http://dx.doi.org/10.1038/s41598-019-42294-8	Intervention	Not a named intervention. Also no radiologist and population not reported
Berbaum KS, Krupinski EA, Schartz KM, Caldwell RT, Madsen MT, Hur S*, et al.* The influence of a vocalized checklist on detection of multiple abnormalities in chest radiography. *Acad Radiol* 2016;**23(4)**:413-20. http://dx.doi.org/10.1016/j.acra.2015.12.017	Intervention	Not a named intervention. Also simulated nodules on CXRs
Cha MJ, Chung MJ, Lee JH, Lee KS. Performance of Deep Learning Model in Detecting Operable Lung Cancer With Chest Radiographs. *J Thorac Imaging* 2019;**34**(2):86-91. http://dx.doi.org/10.1097/RTI.0000000000000388	Intervention	Software not stated; CXR from database, population details and referral not clear
Chen B, Li J, Guo X, Lu G. DualCheXNet: dual asymmetric feature learning for thoracic disease classification in chest X-rays. *Biomed Signal Process Control* 2019;**53 (no pagination)**. http://dx.doi.org/10.1016/j.bspc.2019.04.031	Intervention	Not a named intervention (DualCheXNet), CXRs from a database but referral route of participants not known, validation study for the algorithm
Chen S, Han Y, Lin J, Zhao X, Kong P. Pulmonary nodule detection on chest	Intervention	Not a specified software, paper on development of software

radiographs using balanced convolutional neural network and classic candidate detection. *Artif Intell Med* 2020;**107**:101881. http://dx.doi.org/10.1016/j.artmed.2020.101881		
Chen S, Yao L, Chen B. A parameterized logarithmic image processing method with Laplacian of Gaussian filtering for lung nodule enhancement in chest radiographs. *Med Biol Eng Comput* 2016;**54**(11):1793-806.	Intervention	Not a named intervention. CXRs from a database, referral route not reported, no radiologist input reported, also investigated another non named intervention, validation study for the algorithm
Chetan MR, Dowson N, Price NW, Ather S, Nicolson A, Gleeson FV. Developing an understanding of artificial intelligence lung nodule risk prediction using insights from the Brock model. *Eur Radiol* 2022;**32(8)**:5330-8. http://dx.doi.org/10.1007/s00330-022-08635-4	Intervention	CT scan AI, also screening population
Cho Y, Kim YG, Lee SM, Seo JB, Kim N. Reproducibility of abnormality detection on chest radiographs using convolutional neural network in paired radiographs obtained within a short-term interval. *Sci Rep* 2020;**10**(1):17417. http://dx.doi.org/10.1038/s41598-020-74626-4	Intervention	Not a named intervention
Choi S, Lee O, Kim M. The cut-off values for auto-detection of lung cancer in chest radiography: An example using an unsupervised method. *Biomedical Engineering - Applications, Basis and Communications* 2012;**24(6)**:525-36. http://dx.doi.org/10.4015/S1016237212500482	Intervention	Not a named intervention ('Principle Component Analysis' and 'Texture Information Analysis'), referral route of participants not known
Choi SY, Park S, Kim M, Park J, Choi YR, Jin KN. Evaluation of a deep learning-based computer-aided detection algorithm on chest radiographs: Case-control study. *Medicine* 2021;**100**(16):e25663. http://dx.doi.org/10.1097/MD.0000000000025663	Population	Software eligible. Population referral route not reported but included health screening unit or oncology unit with normal CXRs and not described where those with nodules came from except 'consecutive' cases which were abnormal cases with localized consolidation selected from subjects who visited the emergency department or respiratory medicine
De Boo DW, van Hoorn F, van Schuppen J, Schijf L, Scheerder MJ, Freling NJ*, et al.* Observer training for computer-aided detection of pulmonary nodules in chest radiography. *Eur Radiol* 2012;**22**(8):1659-64. http://dx.doi.org/10.1007/s00330-012-2412-7	Intervention	Software: IQQA Chest, EDDA Technology, Princeton Junction, NJ, USA
Dellios N, Teichgraeber U, Chelaru R, Malich A, Papageorgiou IE. Computer-aided Detection Fidelity of Pulmonary Nodules in Chest Radiograph. *J Clin Imaging Sci* 2017;**7**:8. http://dx.doi.org/10.4103/jcis.JCIS_75_16	Intervention	Riverain manufacturer, but software includes SoftView (bone suppression imaging) and OnGuard (nodule detection) possible version of ClearRead Detect; population with known pulmonary lesions
Dissez G, Tay N, Dyer T, Tam M, Dittrich R, Doyne D*, et al. Enhancing Early Lung Cancer Detection on Chest Radiographs with AI-assistance: A Multi-Reader Study [preprint]*. arXiv.org; 2022. URL:	Population	Software eligible. Was AI+clinician vs clinician. Population from retrospective CXRs collected in one hospital during 2020, referral route not reported

https://arxiv.org/ftp/arxiv/papers/2208/2208.14742.pdf (Accessed 3 January 2023).		
Do Q, Seo W, Shin CW. Automatic algorithm for determining bone and soft-tissue factors in dual-energy subtraction chest radiography. *Biomed Signal Process Control* 2023;**Part 2. 80 (no pagination)**. http://dx.doi.org/10.1016/j.bspc.2022.104354	Intervention	Not a named intervention, also simulation study
Dorri Giv M, Haghighi Borujeini M, Seifi Makrani D, Dastranj L, Yadollahi M, Semyari S*, et al.* Lung Segmentation using Active Shape Model to Detect the Disease from Chest Radiography. *Journal of Biomedical Physics & Engineering* 2021;**11**(6):747-56. http://dx.doi.org/10.31661/jbpe.v0i0.2105-1346	Intervention	Not a specified software, query AI, population from database no details
Dyer T, Dillard L, Harrison M, Morgan TN, Tappouni R, Malik Q*, et al.* Diagnosis of normal chest radiographs using an autonomous deep-learning algorithm. *Clin Radiol* 2021;**76**(6):473.e9-.e15. http://dx.doi.org/10.1016/j.crad.2021.01.015	Population	Not a named intervention (authors employed by behold.ai') but population also not clear (includes A&E, GP, outpatient)
Dyer T, Smith J, Dissez G, Tay N, Malik Q, Morgan TN*, et al. Robustness of an Artificial Intelligence Solution for Diagnosis of Normal Chest X-Rays [preprint]*. arXiv.org; 2022. URL: https://arxiv.org/ftp/arxiv/papers/2209/2209.09204.pdf (Accessed 3 January 2023).	Intervention	Software eligible but stand-alone AI. Study was AI to find normal CXRs. CXRs from retrospective collection and chosen to represent a diverse dataset of NHS patients and care settings. AI vs clinician. No relevant outcomes
Endo K, Kaneko A, Horiuchi Y, Kasuga N, Ishizaki U, Sakai S. Detectability of pulmonary nodules on chest radiographs: bone suppression versus standard technique with single versus dual monitors for visualization. *Japanese Journal of Radiology* 2020;**38**(7):676-82. http://dx.doi.org/10.1007/s11604-020-00952-2	Intervention	Bone suppression imaging using software developed with a deep-learning pattern recognition algorithm
Fischer G, De Silvestro A, Muller M, Frauenfelder T, Martini K. Computer-Aided Detection of Seven Chest Pathologies on Standard Posteroanterior Chest X-Rays Compared to Radiologists Reading Dual-Energy Subtracted Radiographs. *Acad Radiol* 2022;**29**(8):e139-e48. http://dx.doi.org/10.1016/j.acra.2021.09.016	Population	Inpatients and outpatients, most had CXR pre-surgery so unlikely 90% were referred (and Intervention not an adjunct to clinician)
Ghali R, Akhloufi MA. ARSeg: An Attention RegSeg Architecture for CXR Lung Segmentation. Paper presented at: 2022 IEEE 23rd International Conference on Information Reuse and Integration for Data Science (IRI); San Diego, CA, USA. URL: https://doi.org/10.1109/IRI54793.2022.00068	Intervention	Not eligible software
Govindarajan A, Govindarajan A, Tanamala S, Chattoraj S, Reddy B, Agrawal R*, et al.* Role of an Automated Deep Learning Algorithm for Reliable Screening of Abnormality in Chest Radiographs: A Prospective Multicenter Quality Improvement Study. *Diagnostics* 2022;**12**(11):07. http://dx.doi.org/10.3390/diagnostics12112724	Population	qXR but stand alone, population age > 6 years, subgroup results reported but only for normal/abnormal (not nodule), referral status unknown, states 'routine screening'
Guo W, Li Q, Boyce SJ, McAdams HP, Shiraishi J, Doi K*, et al.* A computerized scheme for lung nodule detection in multiprojection chest radiography. *Med*	Intervention	Software not specified, population from database no details

Phys 2012;**39**(4):2001-12. http://dx.doi.org/10.1118/1.3694096		
Hao R, Qiang Z, Qiang Y, Ge L, Zhao J. Automatic diagnosis of pulmonary nodules using a hierarchical extreme learning machine model. *Int J Bio-Inspired Comput* 2018;**11**(3):192–201. http://dx.doi.org/10.1504/ijbic.2018.091748	British library not available	
Homayounieh F, Digumarthy S, Ebrahimian S, Rueckel J, Hoppe BF, Sabel BO*, et al.* An Artificial Intelligence-Based Chest X-ray Model on Human Nodule Detection Accuracy From a Multicenter Study. *JAMA Network Open* 2021;**4**(12):e2141096. http://dx.doi.org/10.1001/jamanetworkopen.2021.41096	Population	Software eligible. Population referral route not reported (retrospective). Was AI+clinician vs clinician
Htike ZZ, Naing WYN, Win SL, Khan S. Computer-Aided Diagnosis of Pulmonary Nodules from Chest X-Rays Using Rotation Forest. Paper presented at: Proceedings of the 2014 International Conference on Computer and Communication Engineering. URL: https://doi.org/10.1109/ICCCE.2014.38	Intervention	Not a named intervention ('proposed system'), CXRs from a database so referral route of participants not known
Huang X, Fang Y, Lu M, Yan F, Yang J, Xu Y. Dual-ray net: Automatic diagnosis of thoracic diseases using frontal and lateral chest x-rays. *Journal of Medical Imaging and Health Informatics* 2020;**10(2)**:348-55. http://dx.doi.org/10.1166/jmihi.2020.2901	Intervention	Not a named intervention, also CXRs from two databases but referral route of participants not known, validation study for the algorithm
Hwang EJ, Park S, Jin KN, Kim JI, Choi SY, Lee JH*, et al.* Development and Validation of a Deep Learning-Based Automated Detection Algorithm for Major Thoracic Diseases on Chest Radiographs. *JAMA Network Open* 2019;**2**(3):e191095. http://dx.doi.org/10.1001/jamanetworkopen.2019.1095	Intervention	Software not stated; CXR datasets from 4 hospitals, referral and details unclear
Jang S, Song H, Shin YJ, Kim J, Lee KW, Lee SS*, et al.* Deep Learning-based Automatic Detection Algorithm for Reducing Overlooked Lung Cancers on Chest Radiographs. *Radiology* 2020;**296(3)**:652-61. http://dx.doi.org/10.1148/radiol.2020200165	Population	Eligible software (Lunit Insight); Population referral status unclear, people with lung cancer and cancer visible on CXR prior to diagnosis, unclear if symptomatic or incidental, control group normal CXR, reason for CXR unclear. AI+clinician vs AI
Jin KN, Kim EY, Kim YJ, Lee GP, Kim H, Oh S*, et al.* Diagnostic effect of artificial intelligence solution for referable thoracic abnormalities on chest radiography: a multicenter respiratory outpatient diagnostic cohort study. *Eur Radiol* 2022;**32**(5):3469-79. http://dx.doi.org/10.1007/s00330-021-08397-5	Population	Software eligible. Population those seen in respiratory outpatients for any lung diseases, nodule/cancer included, no details of referral route and results include any identifiable lesion (nodules or masses, lung consolidation, and pneumothorax) not nodule/cancer specifically
Kao EF, Liu GC, Lee LY, Tsai HY, Jaw TS. Computer-aided detection system for chest radiography: Reducing report turnaround times of examinations with abnormalities. *Acta Radiol* 2015;**56(6)**:696-701. http://dx.doi.org/10.1177/0284185114538017	Intervention	Not a named intervention ('homemade CAD'), referral route of participants not known
Kaviani P, Digumarthy SR, Bizzo BC, Reddy B, Tadepalli M, Putha P*, et al.* Performance of a Chest Radiography AI Algorithm for Detection of Missed or	Population	Population not described, CXRs taken from a database and no information that these would be primary care

Mislabeled Findings: A Multicenter Study. *Diagnostics* 2022;**12**(9):28. http://dx.doi.org/10.3390/diagnostics12092086		referrals, intervention not an adjunct to clinician
Kaviani P, Kalra MK, Digumarthy SR, Gupta RV, Dasegowda G, Jagirdar A*, et al.* Frequency of Missed Findings on Chest Radiographs (CXRs) in an International, Multicenter Study: Application of AI to Reduce Missed Findings. *Diagnostics* 2022;**12**(10):30. http://dx.doi.org/10.3390/diagnostics12102382	Population	Software eligible (Qure.ai). Population not described, not clear if referred with symptoms but only those with 'normal' CXRs were used, intervention not an adjunct to clinician
KCT. *A single-center, randomized, crossover and retrospective pivotal trial to evaluate the efficacy of VUNO Med - Chest X-ray in screening of abnormalities on chest radiograph*. WHO ICTRP; 2019. URL: https://trialsearch.who.int/Trial2.aspx?TrialID=KCT0004147 (Accessed 20 January 2023).	Population	Software eligible. Screening population. Ongoing study no results
KCT. *Diagnosis of lung nodule and lung cancer on screening chest radiographs: comparative clinical trial for evaluation of artificial intelligence-integrated PACS versus conventional PACS*. WHO ICTRP; 2020. URL: https://trialsearch.who.int/Trial2.aspx?TrialID=KCT0005051 (Accessed 20 January 2023).	Population	Software eligible. Screening population (those with respiratory symptoms when CXR performed excluded). Ongoing study no results
Ke Q, Zhang J, Wei W, Połap D, Woźniak M, Kośmider L*, et al.* A neuro-heuristic approach for recognition of lung diseases from X-ray images. *Expert Syst Appl* 2019;**126**(C):218–32. http://dx.doi.org/10.1016/j.eswa.2019.01.060	Intervention	Not a named intervention, CXRs from 3 databases but referral route of participants not known, validation study for the algorithm
Kim EY, Kim YJ, Choi WJ, Jeon JS, Kim MY, Oh DH*, et al.* Concordance rate of radiologists and a commercialized deep-learning solution for chest X-ray: Real-world experience with a multicenter health screening cohort. *PLoS ONE [Electronic Resource]* 2022;**17**(2):e0264383. http://dx.doi.org/10.1371/journal.pone.0264383	Population	Health screening population, no description of referral route or reason for CXR other than for a screening test, outcomes were broad groups of thoracic abnormalities (inactive, insignificant abnormal, and significant abnormal lesions)
Kim H, Park CM, Goo JM. Test-retest reproducibility of a deep learning-based automatic detection algorithm for the chest radiograph. *Eur Radiol* 2020;**30**(4):2346-55. http://dx.doi.org/10.1007/s00330-019-06589-8	Population	Eligible software, Population undergoing pre-op CXR, comparator not eligible
Kligerman S, Cai L, White CS. The effect of computer-aided detection on radiologist performance in the detection of lung cancers previously missed on a chest radiograph. *J Thorac Imaging* 2013;**28**(4):244-52. http://dx.doi.org/10.1097/RTI.0b013e31826c29ec	Intervention	OnGuard (Riverain)+ radiologist vs radiologist alone, population were lung cancer previously missed on CXR (CXR from before diagnosis) - referral status unknown, unclear if incidental or symptomatic
Koo YH, Shin KE, Park JS, Lee JW, Byun S, Lee H. Extravalidation and reproducibility results of a commercial deep learning-based automatic detection algorithm for pulmonary nodules on chest radiographs at tertiary hospital. *J Med Imaging Radiat Oncol* 2021;**65**(1):15-22. http://dx.doi.org/10.1111/1754-9485.13105	Population	Software eligible. Population referral route not reported but from a tertiary centre and included CXRs with known nodules and without, the prevalence of nodules was 46.5%, includes AI+clinician vs AI alone

Laksshmi KSG, Umagandhi R. False Positive Reduction Based on Anatomical Characterization Using Deep Learning Neural Network in Lung Nodule Detection. *European Journal of Molecular and Clinical Medicine* 2020;**7(8)**:5296-303.	Intervention	Not eligible software
Lee KH, Goo JM, Park CM, Lee HJ, Jin KN. Computer-aided detection of malignant lung nodules on chest radiographs: effect on observers' performance. *Korean Journal of Radiology* 2012;**13**(5):564-71. http://dx.doi.org/10.3348/kjr.2012.13.5.564	Intervention	CAD: IQQA-Chest, EDDA Technology, Princeton Junction, NJ, USA; population: retrospective selection of malignant nodules and normal cases
Lee YW, Huang SK, Chang RF. CheXGAT: A disease correlation-aware network for thorax disease diagnosis from chest X-ray images. *Artif Intell Med* 2022;**132 (no pagination)**. http://dx.doi.org/10.1016/j.artmed.2022.102382	Intervention	Not eligible software
Li F, Engelmann R, Armato SG, 3rd, MacMahon H. Computer-aided nodule detection system: results in an unselected series of consecutive chest radiographs. *Acad Radiol* 2015;**22**(4):475-80. http://dx.doi.org/10.1016/j.acra.2014.11.008	Intervention	ClearRead Detect (eligible) and SoftView v2.4 (bone suppression imaging), not with radiologist; population unclear - had CT on same day, outcomes for nodule detection but comparator is radiologist+CXR+CT
Li X, Shen L, Luo S. A Solitary Feature-Based Lung Nodule Detection Approach for Chest X-Ray Radiographs. *IEEE Journal of Biomedical & Health Informatics* 2018;**22**(2):516-24. http://dx.doi.org/10.1109/JBHI.2017.2661805	Intervention	Not a named intervention. CXRs from two databases and referral route not reported, no radiologist reported, validation study of the algorithm
Li X, Shen L, Xie X, Huang S, Xie Z, Hong X*, et al.* Multi-resolution convolutional networks for chest X-ray radiograph based lung nodule detection. *Artif Intell Med* 2020;**103**:101744. http://dx.doi.org/10.1016/j.artmed.2019.101744	Intervention	Population from databases, no details of referral; software not named
Liang CH, Liu YC, Wu MT, Garcia-Castro F, Alberich-Bayarri A, Wu FZ. Identifying pulmonary nodules or masses on chest radiography using deep learning: external validation and strategies to improve clinical practice. *Clin Radiol* 2020;**75**(1):38-45. http://dx.doi.org/10.1016/j.crad.2019.08.005	Intervention	Population referral unclear, software not eligible (QUIBIM)
Liu H, Wang L, Nan Y, Jin F, Wang Q, Pu J. SDFN: Segmentation-based deep fusion network for thoracic disease classification in chest X-ray images. *Comput Med Imaging Graph* 2019;**75**:66-73. http://dx.doi.org/10.1016/j.compmedimag.2019.05.005	Intervention	From known SR not in Endnote. Not a named intervention (DenseNet), CXRs from databases, referral route of participants not known, validation study for the algorithm
Louati H, Louati A, Bechikh S, Said LB. Design and Compression Study for Convolutional Neural Networks Based on Evolutionary Optimization for Thoracic X-Ray Image Classification. Paper presented at: Computational Collective Intelligence: 14th International Conference, ICCCI 2022, Hammamet, Tunisia, September 28–30, 2022, Proceedings; Hammamet, Tunisia. URL: https://doi.org/10.1007/978-3-031-16014-1_23	Intervention	Not a named intervention, also no details of where CXRs were from, validation study for the algorithm
Majkowska A, Mittal S, Steiner DF, Reicher JJ, McKinney SM, Duggan GE*, et al.* Chest radiograph interpretation with deep learning models: Assessment with	Intervention	Not a named intervention. Referral route unclear but one database consecutive inpatient and outpatient images and

radiologist-adjudicated reference standards and population-adjusted evaluation. *Radiology* 2020;**294(2)**:421-31.		the other all CXRS from multiple different hospitals
Malik H, Anees T, Mui Zzud D. BDCNet: multi-classification convolutional neural network model for classification of COVID-19, pneumonia, and lung cancer from chest radiographs. *Multimedia Systems* 2022;**28**(3):815-29. http://dx.doi.org/10.1007/s00530-021-00878-3	Intervention	Compares 4 named ineligible software
Martinez-Machado E, Perez-Diaz M, Orozco-Morales R. Automated System for the Detection of Lung Nodules. Paper presented at: Progress in Artificial Intelligence and Pattern Recognition: 7th International Workshop on Artificial Intelligence and Pattern Recognition, IWAIPR 2021, Havana, Cuba, October 5–7, 2021, Proceedings; Havana, Cuba. URL: https://doi.org/10.1007/978-3-030-89691-1_33	Intervention	Not a named intervention. Retrospective database of CXRs with no discussion of referral route of pts, and unclear if radiologist input, validation study of the algorithm
Mathew TE, Sugelanandh O. Lung Cancer Classification Using Extreme-Anfiswith Red Fox Optimization Algorithm. *Neuroquantology* 2022;**20(6)**:1839-46. http://dx.doi.org/10.14704/nq.2022.20.6.NQ22183	Intervention	Not a named intervention, also no details of where CXRs were from, validation study for the algorithm
Mazzone PJ, Obuchowski N, Phillips M, Risius B, Bazerbashi B, Meziane M. Lung cancer screening with computer aided detection chest radiography: design and results of a randomized, controlled trial. *PLoS ONE [Electronic Resource]* 2013;**8**(3):e59650. http://dx.doi.org/10.1371/journal.pone.0059650	Intervention	OnGuard 5.0 (Riverain), vs placebo CXR (RCT), screening of a high-risk population
Meraj T, Rauf HT, Zahoor S, Hassan A, Lali MI, Ali L*, et al.* Lung nodules detection using semantic segmentation and classification with optimal features. *Neural Comput Appl* 2021;**33**(17):10737–50. http://dx.doi.org/10.1007/s00521-020-04870-2	Intervention	CT images
Messerli M, Kluckert T, Knitel M, Rengier F, Warschkow R, Alkadhi H*, et al.* Computer-aided detection (CAD) of solid pulmonary nodules in chest x-ray equivalent ultralow dose chest CT - first in-vivo results at dose levels of 0.13mSv. *Eur J Radiol* 2016;**85**(12):2217-24. http://dx.doi.org/10.1016/j.ejrad.2016.10.006	Intervention	Not CXR
Meziane M, Mazzone P, Novak E, Lieber ML, Lababede O, Phillips M*, et al.* A comparison of four versions of a computer-aided detection system for pulmonary nodules on chest radiographs. *J Thorac Imaging* 2012;**27**(1):58-64. http://dx.doi.org/10.1097/RTI.0b013e3181f240bc	Intervention	Compares 4 generations of CAD software: RapidScreen1.1andOnGuard3.0, 4.0, and5.0 (RiverainMedical)
Miyoshi T, Yoshida J, Aramaki N, Matsumura Y, Aokage K, Hishida T*, et al.* Effectiveness of Bone Suppression Imaging in the Detection of Lung Nodules on Chest Radiographs: Relevance to Anatomic Location and Observer's Experience. *J Thorac Imaging* 2017;**32**(6):398-405. http://dx.doi.org/10.1097/RTI.0000000000000299	Intervention	Not a named intervention and not AI. Also referral route of participants CXRs not reported

Multicenter, multi-reader, multicase (MRMC) study on the performance of AI for pulmonary nodule detection on chest radiographs with accompanying chest CT for ground trothing [Ongoing study. AIC data from CS]. In.,	Population	Software eligible. AIC data from CS. Ongoing study. Was AI+clinician vs clinician. Referral route of CXRs unclear were from academic health centres with CT on the same day
Nam JG, Hwang EJ, Kim DS, Yoo SJ, Choi H, Goo JM*, et al.* Undetected Lung Cancer at Posteroanterior Chest Radiography: Potential Role of a Deep Learning-based Detection Algorithm. *Radiology Cardiothoracic Imaging* 2020;**2**(6):e190222. http://dx.doi.org/10.1148/ryct.2020190222	Population	Software eligible. CXRs from people with confirmed lung CA initially undetected on CXR - unclear referral route or if CXR for symptoms or no symptoms, also unclear where the 'normal' x-rays are from, is algorithm + radiologist vs radiologist
Nam JG, Kim M, Park J, Hwang EJ, Lee JH, Hong JH*, et al.* Development and validation of a deep learning algorithm detecting 10 common abnormalities on chest radiographs. *Eur Respir J* 2021;**57**(5). http://dx.doi.org/10.1183/13993003.03061-2020	Population	Software named as DLAD-10, company submission states is INSIGHT. CXRs from retrospective databases, referral route unknown, but had CT on the same day. Also simulation validation from CXRs from emergency department CXRs. Intervention looking at 10 different lung conditions. Only the simulation validation set was AI+clinician vs clinician.
Nam JG, Park S, Hwang EJ, Lee JH, Jin KN, Lim KY*, et al.* Development and Validation of Deep Learning-based Automatic Detection Algorithm for Malignant Pulmonary Nodules on Chest Radiographs. *Radiology* 2019;**290**(1):218-28. http://dx.doi.org/10.1148/radiol.2018180237	Intervention	Funded by Lunit but software not stated; retrospective data set, population and referral unclear, comparison not eligible (for nodule detection observers asked if decision changed with results of software)
Napoleon D, Kalaiarasi I. Classifying Lung Cancer as Benign and Malignant Nodule Using ANN of Back-Propagation Algorithm and GLCM Feature Extraction on Chest X-Ray Images. *Wirel Pers Commun* 2022;**126**(1):167–95. http://dx.doi.org/10.1007/s11277-022-09594-1	Intervention	Not a named intervention, also CXRs with known abnormalities to distinguish between malignant and benign with no details of referral route
Nasser AA, Akhloufi MA. Chest Diseases Classification Using CXR and Deep Ensemble Learning. Paper presented at: Proceedings of the 19th International Conference on Content-based Multimedia Indexing; Graz, Austria. URL: https://doi.org/10.1145/3549555.3549581).	Intervention	Not a named intervention, also not lung cancer or nodules
Nasser AA, Akhloufi MA. Classification of CXR Chest Diseases by Ensembling Deep Learning Models. Paper presented at: 2022 IEEE 23rd International Conference on Information Reuse and Integration for Data Science (IRI); San Diego, CA, USA. URL: https://doi.org/10.1109/IRI54793.2022.00062	Intervention	Not a named intervention, also not lung cancer or nodules
NCT. *xrAI - Improving Quality and Efficiency in Chest Radiograph Interpretation*. ClinicalTrials.gov; 2019. URL: https://clinicaltrials.gov/show/NCT04153045 (Accessed 20 January 2023).	Intervention	Not a named intervention, study completed but no results posted
NCT. *xrAI - Improving Quality and Efficiency in Chest Radiograph Interpretation by Radiologists*. ClinicalTrials.gov; 2020. URL:	Duplicate	Duplicate

https://clinicaltrials.gov/show/NCT04221100 (Accessed 20 January 2023).		
Novak RD, Novak NJ, Gilkeson R, Mansoori B, Aandal GE. A comparison of computer-aided detection (CAD) effectiveness in pulmonary nodule identification using different methods of bone suppression in chest radiographs. *J Digit Imaging* 2013;**26**(4):651-6. http://dx.doi.org/10.1007/s10278-012-9565-4	Intervention	ClearRead Detect (eligible) but unclear if with radiologist, comparison is other CAD image types; Patients with pulmonary nodules confirmed by CT or pathology-proven CA selected, and negative cases selected, referral status of all unclear
Park S, Lee SM, Lee KH, Jung KH, Bae W, Choe J*, et al.* Deep learning-based detection system for multiclass lesions on chest radiographs: comparison with observer readings. *Eur Radiol* 2020;**30**(3):1359-68. http://dx.doi.org/10.1007/s00330-019-06532-x	Intervention	Unclear if referred from primary care, software not named and not as adjunct
Pesce E, Joseph Withey S, Ypsilantis PP, Bakewell R, Goh V, Montana G. Learning to detect chest radiographs containing pulmonary lesions using visual attention networks. *Med Image Anal* 2019;**53**:26-38. http://dx.doi.org/10.1016/j.media.2018.12.007	Intervention	Software not stated; population unclear
Peters AA, Decasper A, Munz J, Klaus J, Loebelenz LI, Hoffner MKM*, et al.* Performance of an AI based CAD system in solid lung nodule detection on chest phantom radiographs compared to radiology residents and fellow radiologists. *J Thorac Dis* 2021;**13**(5):2728-37. http://dx.doi.org/10.21037/jtd-20-3522	Population	Simulation study
Pham HH, Le TT, Tran DQ, Ngo DT, Nguyen HQ. Interpreting chest X-rays via CNNs that exploit hierarchical disease dependencies and uncertainty labels. *Neurocomputing* 2021;**437**:186-94. http://dx.doi.org/10.1016/j.neucom.2020.03.127	Intervention	Not a named intervention, also CXRs from a database but referral route of participants not known, validation study for the algorithm
Pooch EHP, Alva TAP, Becker CDL. A Deep Learning Approach for Pulmonary Lesion Identification in Chest Radiographs. Paper presented at: Intelligent Systems: 9th Brazilian Conference, BRACIS 2020, Rio Grande, Brazil, October 20–23, 2020, Proceedings, Part I; Rio Grande, Brazil. URL: https://doi.org/10.1007/978-3-030-61377-8_14	Intervention	Not a named intervention, also CXRs from a database but referral route of participants not known, validation study for the algorithm
Putha P, Tadepalli M, Reddy B, Raj T, Chiramal JA, Govil S*, et al. Can artificial intelligence reliably report chest x-rays?: Radiologist validation of an algorithm trained on 2.3 million x-rays [Preprint].* arXiv.org; 2018. URL: https://arxiv.org/pdf/1807.07455.pdf (Accessed 4 January 2023).	Intervention	Software not named, CXRs from databases including participants from inpatient and outpatient and no route of referral known, validation study for the algorithm to identify numerous abnormalities, radiologist vs AI only for nodules
Rajagopalan K, Babu S. The detection of lung cancer using massive artificial neural network based on soft tissue technique. *BMC Med Inform Decis Mak* 2020;**20**(1):282. http://dx.doi.org/10.1186/s12911-020-01220-z	Intervention	Not a named intervention
Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H*, et al.* Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt	Intervention	CheXNeXT, training and validation study using dataset (ChestX-ray-14)no details

algorithm to practicing radiologists. *PLoS Medicine / Public Library of Science* 2018;**15**(11):e1002686. http://dx.doi.org/10.1371/journal.pmed.1002686		
Ridder K, Preuhs A, Mertins A, Joerger C. *Routine Usage of AI-based Chest X-ray Reading Support in a Multi-site Medical Supply Center [Preprint]*. arXiv.org; 2022. URL: https://arxiv.org/ftp/arxiv/papers/2210/2210.10779.pdf (Accessed 3 January 2023).	Intervention	Software eligible but standalone. No details of where CXRs were from, abstract only, no comparator
Saba T. Automated lung nodule detection and classification based on multiple classifiers voting. *Microsc Res Tech* 2019;**82**(9):1601-9. http://dx.doi.org/10.1002/jemt.23326	Intervention	Not CXR (CT)
Samei E, Majdi-Nasab N, Dobbins JT, 3rd, McAdams HP. Biplane correlation imaging: a feasibility study based on phantom and human data. *J Digit Imaging* 2012;**25**(1):137-47. http://dx.doi.org/10.1007/s10278-011-9392-z	Intervention	Not a named intervention, simulated cases and some human cases but no details of where from
Schalekamp S, van Ginneken B, Heggelman B, Imhof-Tas M, Somers I, Brink M*, et al.* New methods for using computer-aided detection information for the detection of lung nodules on chest radiographs. *Br J Radiol* 2014;**87**(1036):20140015. http://dx.doi.org/10.1259/bjr.20140015	Population	Intervention: ClearRead Detect with ClearRead Bone Suppression + radiologist; same readers for intervention and comparator; CXR retrospectively selected, derived from clinically indicated examinations, referral route unclear
Schalekamp S, van Ginneken B, Koedam E, Snoeren MM, Tiehuis AM, Wittenberg R*, et al.* Computer-aided detection improves detection of pulmonary nodules in chest radiographs beyond the support by bone-suppressed images. *Radiology* 2014;**272**(1):252-61. http://dx.doi.org/10.1148/radiol.14131315	Population	ClearRead Detect with ClearRead Bone Suppression + radiologist; same readers for intervention and comparator; referral route unclear: patients retrospectively selected with presence of a solid solitary lesion and CT within 3 months, control patients negative CXR and CT within 6 months
Seyyed-Kalantari L, Liu G, McDermott M, Chen IY, Ghassemi M. CheXclusion: Fairness gaps in deep chest X-ray classifiers. *Pac Symp Biocomput* 2021;**26**:232-43. https://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=med19&AN=33691020	Intervention	Not a named intervention
Shi Z, Ma J, Feng Y, He L, Suzuki K. Evaluation of MTANNs for eliminating false-positive with different computer aided pulmonary nodules detection software. *Pak J Pharm Sci* 2015;**28**(6 Suppl):2311-6.	Intervention	Additional algorithm applied to named interventions but on simulations
Shi Z, Xu B, Zhao M, Zhao J, Wang Y, Liu Y*, et al.* A joint ROI extraction filter for computer aided lung nodule detection. *Biomed Mater Eng* 2015;**26(Supplement 1)**:S1491-S9. http://dx.doi.org/10.3233/BME-151448	Intervention	CT scan
Shimazaki A, Ueda D, Choppin A, Yamamoto A, Honjo T, Shimahara Y*, et al.* Deep learning-based algorithm for lung cancer detection on chest radiographs using the segmentation method. *Sci Rep* 2022;**12**(1):727.	Intervention	Not a named interventions, also confirmed lung cancer cases only and not clear when CXRs performed

http://dx.doi.org/10.1038/s41598-021-04667-w		
Sim Y, Chung MJ, Kotter E, Yune S, Kim M, Do S*, et al.* Deep Convolutional Neural Network-based Software Improves Radiologist Detection of Malignant Lung Nodules on Chest Radiographs. *Radiology* 2020;**294**(1):199-209. http://dx.doi.org/10.1148/radiol.2019182465	Population	Eligible software (and with radiologist on re-read); Population includes normal CXR from health screening population and CXR with lung cancer at tertiary hospital - referral unclear
Singh A, Lall B, Panigrahi BK, Agrawal A, Thangakunam B, Christopher DJ. Deep LF-Net: Semantic lung segmentation from Indian chest radiographs including severely unhealthy images. *Biomed Signal Process Control* 2021;**68 (no pagination)**. http://dx.doi.org/10.1016/j.bspc.2021.102666	Intervention	Not eligible software, population from three datasets, referral status unknown
Singh R, Kalra MK, Nitiwarangkul C, Patti JA, Homayounieh F, Padole A*, et al.* Deep learning in chest radiography: Detection of findings and presence of change. *PLoS ONE [Electronic Resource]* 2018;**13**(10):e0204155. http://dx.doi.org/10.1371/journal.pone.0204155	Intervention	Employees of Qure.ai and software referred to as Qure AI, appears to be stand alone; population from CHestX-ray8 datatset, no details; outcome not nodules or cancer
Sirshar M, Hassan T, Akram MU, Khan SA. An incremental learning approach to automatically recognize pulmonary diseases from the multi-vendor chest radiographs. *Comput Biol Med* 2021;**134**(C):9. http://dx.doi.org/10.1016/j.compbiomed.2021.104435	Intervention	Not a named intervention, also CXRs from various databases but referral route of participants not known, validation study for the algorithm
Stubblefield JW, Cooksey L, Causey J, Qualls J, Bellis E, Ashby C*, et al. Artificial Intelligence Algorithms for Medical Imaging and Healthcare*: Arkansas State University; 2021.	Study design	Thesis, full text not retrieved
Szucs-Farkas Z, Schick A, Cullmann JL, Ebner L, Megyeri B, Vock P*, et al.* Comparison of dual-energy subtraction and electronic bone suppression combined with computer-aided detection on chest radiographs: effect on human observers' performance in nodule detection. *AJR American Journal of Roentgenology* 2013;**200**(5):1006-13. http://dx.doi.org/10.2214/AJR.12.8877	Intervention	Riverain manufacturer, but software includes SoftView (bone suppression imaging) and OnGuard (nodule detection) – query early version of ClearRead Detect; population retrospectively selected with pulmonary nodules
Tam M, Dyer T, Dissez G, Morgan TN, Hughes M, Illes J*, et al.* Augmenting lung cancer diagnosis on chest radiographs: positioning artificial intelligence to improve radiologist performance. *Clin Radiol* 2021;**76**(8):607-14. http://dx.doi.org/10.1016/j.crad.2021.03.021	Population	Software eligible. Population referral route not reported. Includes CXRs with difficult to locate nodules and CXRs with no nodules. includes AI+clinician vs AI alone but is simulating what might happen if the AI alone was used as triage
Teng PH, Liang CH, Lin Y, Alberich-Bayarri A, Gonzalez RL, Li PW*, et al.* Performance and educational training of radiographers in lung nodule or mass detection: Retrospective comparison with different deep learning algorithms. *Medicine* 2021;**100**(23):e26270. http://dx.doi.org/10.1097/MD.0000000000026270	Intervention	QUIBIM Chest X-ray Classifier (stated in abstract)
Toda N, Hashimoto M, Iwabuchi Y, Nagasaka M, Takeshita R, Yamada M*, et al.*	Intervention	Not a named intervention

Validation of deep learning-based computer-aided detection software use for interpretation of pulmonary abnormalities on chest radiographs and examination of factors that influence readers' performance and final diagnosis. *Japanese Journal of Radiology* 2022;**19**:19. http://dx.doi.org/10.1007/s11604-022-01330-w		
Ueda D, Yamamoto A, Shimazaki A, Walston SL, Matsumoto T, Izumi N*, et al.* Artificial intelligence-supported lung cancer detection by multi-institutional readers with multi-vendor chest radiographs: a retrospective clinical validation study. *BMC Cancer* 2021;**21**(1):1120. http://dx.doi.org/10.1186/s12885-021-08847-9	Intervention	Not a named intervention, also population unclear referral route as retrospectively collected and includes CXRs from confirmed lung cancer cases and those without nodules
van Beek EJR, Ahn JS, Kim MJ, Murchison JT. Validation study of machine-learning chest radiograph software in primary and emergency medicine. *Clin Radiol* 2022;**25**:25. http://dx.doi.org/10.1016/j.crad.2022.08.129	Intervention	Intervention eligible (Lunit INSIGHT CXR (Lunit)) but not an adjunct to clinician, CXRs from referrals from primary care and ED and reported separately, compares AI alone vs human reader alone
Wang H, Jia H, Lu L, Xia Y. Thorax-Net: An Attention Regularized Deep Neural Network for Classification of Thoracic Diseases on Chest Radiography. *IEEE J Biomed Health Inform* 2020;**24**(2):475-85. http://dx.doi.org/10.1109/jbhi.2019.2928369	Intervention	From known SR not in Endnote. Not a named intervention.
Wang Q, Kang W, Wu C, Wang B. Computer-aided detection of lung nodules by SVM based on 3D matrix patterns. *Clin Imaging* 2013;**37**(1):62-9. http://dx.doi.org/10.1016/j.clinimag.2012.02.003	Intervention	Not CXR
Wozniak M, Polap D, Capizzi G, Sciuto GL, Kosmider L, Frankiewicz K. Small lung nodules detection based on local variance analysis and probabilistic neural network. *Comput Methods Programs Biomed* 2018;**161**:173-80. http://dx.doi.org/10.1016/j.cmpb.2018.04.025	Intervention	Software not stated, no details on population
Xu Y, Ma D, He W. Assessing the use of digital radiography and a real-time interactive pulmonary nodule analysis system for large population lung cancer screening. *Eur J Radiol* 2012;**81**(4):e451-6. http://dx.doi.org/10.1016/j.ejrad.2011.04.031	Intervention	Software: IQQA®-Chest Workstation
Yates EJ, Yates LC, Harvey H. Machine learning "red dot": open-source, cloud, deep convolutional neural networks in chest radiograph binary normality classification. *Clin Radiol* 2018;**73**(9):827-31. http://dx.doi.org/10.1016/j.crad.2018.05.015	Intervention	Intervention not relevant, ChestX-ray14 dataset
Yoo H, Kim EY, Kim H, Choi YR, Kim MY, Hwang SH*, et al.* Artificial Intelligence-Based Identification of Normal Chest Radiographs: A Simulation Study in a Multicenter Health Screening Cohort. *Korean Journal of Radiology*	Population	Health "screening" population, aim was to help remove normal CXRs so unlikely referred symptomatic or incidental

2022;**23**(10):1009-18. http://dx.doi.org/10.3348/kjr.2022.0189		
Yoo H, Lee SH, Arru CD, Doda Khera R, Singh R, Siebert S*, et al.* AI-based improvement in lung cancer detection on chest radiographs: results of a multi-reader study in NLST dataset. *Eur Radiol* 2021;**31**(12):9664-74. http://dx.doi.org/10.1007/s00330-021-08074-7	Population	Health screening population from an RCT of lung cancer screening
Zhang Z, Yang J, Zhao J. An automatic detection model of pulmonary nodules based on deep belief network. *Int J Wire Mob Comput* 2019;**16**(1):7–13.	British library not available	
Zheng S, Shen Z, Pei C, Ding W, Lin H, Zheng J*, et al.* Interpretative computer-aided lung cancer diagnosis: From radiology analysis to malignancy evaluation. *Comput Methods Prog Biomed* 2021;**210**(C):11. http://dx.doi.org/10.1016/j.cmpb.2021.106363	Intervention	Not eligible software, R2MNet, for CT not CXR

Non-Cancer:

Reference	**Reason for exclusion**	**Comments**
Adu K, Yu Y, Cai J, Tattrah VD, Ansere JA, Tashi N. S-CCCapsule: Pneumonia detection in chest X-ray images using skip-connected convolutions and capsule neural network. *J Intell Fuzzy Syst* 2021;**41**(1):757–81. http://dx.doi.org/10.3233/jifs-202638	Outcome	Not named commercial software, not AI+clinician, no outcomes of relevance
Afzali A, Babapour Mofrad F, Pouladian M. Contour-based lung shape analysis in order to tuberculosis detection: modeling and feature description. *Med Biol Eng Comput* 2020;**58**(9):1965-86. http://dx.doi.org/10.1007/s11517-020-02192-y	Outcome	Not named commercial software, not AI+clinician, no outcomes of relevance
Albahli S. A deep neural network to distinguish COVID-19 from other chest diseases using X-ray images. *Current Medical Imaging* 2021;**17(1)**:109-19. http://dx.doi.org/10.2174/1573405616666200604163954	Outcome	Not named commercial software, not AI+clinician, no outcomes of relevance although nodules included
Anter AM, Oliva D, Thakare A, Zhang Z. AFCM-LSMA: New intelligent model based on Lévy slime mould algorithm and adaptive fuzzy C-means for identification of COVID-19 infection from chest X-ray images. *Adv Eng Inform* 2021;**49**(C):13. http://dx.doi.org/10.1016/j.aei.2021.101317	Outcome	Not named commercial software, not AI+clinician, no outcomes of relevance
Bharati S, Podder P, Mondal MRH. Hybrid deep learning for detecting lung diseases from X-ray images. *Informatics in Medicine Unlocked* 2020;**20 (no pagination)**. http://dx.doi.org/10.1016/j.imu.2020.100391	Intervention	Not named commercial software, not AI+clinician
Bhardwaj P, Kaur A. A novel and efficient deep learning approach for COVID-19 detection using X-ray imaging modality. *International Journal of Imaging Systems & Technology* 2021;**31**(4):1775-91.	Outcome	Not named commercial software, not AI+clinician, no outcomes of relevance

http://dx.doi.org/10.1002/ima.22627		
Codlin AJ, Dao TP, Vo LNQ, Forse RJ, Van Truong V, Dang HM*, et al.* Independent evaluation of 12 artificial intelligence solutions for the detection of tuberculosis. *Sci Rep* 2021;**11**(1):23895. http://dx.doi.org/10.1038/s41598-021-03265-0	Outcome	AI outcomes were abnormal opacities/cavitation/lesions possibly caused by TB or normal which included abnormal non-tubercular origin
Damania K, Pawar PM, Pramanik R. Convolutional Neural Networks for Detection of COVID-19 From Chest X-Rays. *Int J Ambient Comput Intell* 2022;**13**(1):1–21. http://dx.doi.org/10.4018/ijaci.300793	Outcome	Not named commercial software, not AI+clinician, no outcomes of relevance
El-Bana S, Al-Kabbany A, Sharkas M. A multi-task pipeline with specialized streams for classification and segmentation of infection manifestations in COVID-19 scans. *PeerJ Computer Science* 2020;**6**:e303. http://dx.doi.org/10.7717/peerj-cs.303	Outcome	Simulation study, not commercial named software, no outcomes of relevance
Engle E, Gabrielian A, Long A, Hurt DE, Rosenthal A. Performance of Qure.ai automatic classifiers against a large annotated database of patients with diverse forms of tuberculosis. *PLoS ONE [Electronic Resource]* 2020;**15**(1):e0224445. http://dx.doi.org/10.1371/journal.pone.0224445	Intervention	Incidental population. Retrospective evaluation of CXRs of people with TB, nodules was an outcome, unclear referral route and not AI+clinician
Ezzat D, Hassanien AE, Ella HA. An optimized deep learning architecture for the diagnosis of COVID-19 disease based on gravitational search optimization. *Applied Soft Computing* 2021;**98**:106742. http://dx.doi.org/10.1016/j.asoc.2020.106742	Outcome	Not named commercial software, not AI+clinician, no outcomes of relevance
Fehr J, Konigorski S, Olivier S, Gunda R, Surujdeen A, Gareta D*, et al.* Computer-aided interpretation of chest radiography reveals the spectrum of tuberculosis in rural South Africa. *npj Digital Medicine* 2021;**4(1) (no pagination)**. http://dx.doi.org/10.1038/s41746-021-00471-y	Outcome	Not named commercial software, not AI+clinician, no outcomes of relevance
Gayathri JL, Abraham B, Sujarani MS, Nair MS. A computer-aided diagnosis system for the classification of COVID-19 and non-COVID-19 pneumonia on chest X-ray images by integrating CNN with sparse autoencoder and feed forward neural network. *Comput Biol Med* 2022;**141 (no pagination)**. http://dx.doi.org/10.1016/j.compbiomed.2021.105134	Outcome	Not named commercial software, not AI+clinician, no outcomes of relevance
Gipson J, Tang V, Seah J, Kavnoudias H, Zia A, Lee R*, et al.* Diagnostic accuracy of a commercially available deep-learning algorithm in supine chest radiographs following trauma. *Br J Radiol* 2022;**95**(1134):20210979. http://dx.doi.org/10.1259/bjr.20210979	Outcome	Software eligible. Participants were CXR following 'presenting' with blunt trauma, referral route unclear, not AI+clinician, no lung cancer or nodules
Govindarajan S, Swaminathan R. Analysis of Tuberculosis in Chest Radiographs for Computerized Diagnosis using Bag of Keypoint Features. *J Med Syst* 2019;**43**(4):1–9. http://dx.doi.org/10.1007/s10916-019-1222-8	Outcome	Not named commercial software, not AI+clinician, no outcomes of relevance
Gupta A, Sheth P, Xie P. Neural architecture search for pneumonia diagnosis	Outcome	Not named commercial software, not AI+clinician, no

from chest X-rays. *Sci Rep* 2022;**12**(1):11309. http://dx.doi.org/10.1038/s41598-022-15341-0		outcomes of relevance
Haghanifar A, Majdabadi MM, Choi Y, Deivalakshmi S, Ko S. COVID-CXNet: Detecting COVID-19 in frontal chest X-ray images using deep learning. *Multimedia Tools Appl* 2022;**81**(21):30615–45. http://dx.doi.org/10.1007/s11042-022-12156-z	Outcome	Not named commercial software, not AI+clinician, no outcomes of relevance
Hajjej F, Ayouni S, Hasan M, Abir T, Kaur A. Automatic Detection of Cases of COVID-19 Pneumonia from Chest X-ray Images and Deep Learning Approaches. *Intell Neuroscience* 2022;**2022**:8. http://dx.doi.org/10.1155/2022/7451551	Outcome	Not named commercial software, not AI+clinician, no outcomes of relevance
Hipolito Canario DA, Fromke E, Patetta MA, Eltilib MT, Reyes-Gonzalez JP, Rodriguez GC*, et al.* Using artificial intelligence to risk stratify COVID-19 patients based on chest X-ray findings. *Intelligence-Based Medicine* 2022;**6**:100049. http://dx.doi.org/10.1016/j.ibmed.2022.100049	Outcome	Not commercial software (modified Qxr), not AI+clinician, no outcomes of lung cancer or nodules
Hong W, Hwang EJ, Lee JH, Park J, Goo JM, Park CM. Deep Learning for Detecting Pneumothorax on Chest Radiographs after Needle Biopsy: Clinical Implementation. *Radiology* 2022;**303**(2):433-41. http://dx.doi.org/10.1148/radiol.211706	Outcome	Software eligible (Lunit Insight), AI+clinician vs AI, no nodule or lung cancer outcomes
Hwang EJ, Kim H, Yoon SH, Goo JM, Park CM. Implementation of a Deep Learning-Based Computer-Aided Detection System for the Interpretation of Chest Radiographs in Patients Suspected for COVID-19. *Korean Journal of Radiology* 2020;**21**(10):1150-60. http://dx.doi.org/10.3348/kjr.2020.0536	Outcome	Software eligible (Lunit Insight), AI+clinician vs AI, outcomes presence versus absence of any suggestion of pneumonia. Nodules could be a reason for a false positive - N(%) are reported but this is not against a reference standard /any diagnostic accuracy outcomes
Irmak E. Implementation of convolutional neural network approach for COVID-19 disease detection. *Physiol Genomics* 2020;**52**(12):590-601. http://dx.doi.org/10.1152/physiolgenomics.00084.2020	Outcome	Not named commercial software, not AI+clinician, no outcomes of relevance
Jaeger S, Karargyris A, Candemir S, Folio L, Siegelman J, Callaghan F*, et al.* Automatic tuberculosis screening using chest radiographs. *IEEE Trans Med Imaging* 2014;**33(2)**:233-45. http://dx.doi.org/10.1109/TMI.2013.2284099	Outcome	Not named commercial software, not AI+clinician, no outcomes of relevance
Jangam E, Barreto AAD, Annavarapu CSR. Automatic detection of COVID-19 from chest CT scan and chest X-Rays images using deep learning, transfer learning and stacking. *Applied intelligence (Dordrecht, Netherlands)* 2022;**52**(2):2243-59. http://dx.doi.org/10.1007/s10489-021-02393-4	Outcome	Not named commercial software, not AI+clinician, no outcomes of relevance
Kagujje M, Kerkhoff AD, Nteeni M, Dunn I, Mateyo K, Muyoyeta M. The performance of computer-aided detection digital chest X-ray reading technologies for triage of active Tuberculosis among persons with a history of previous Tuberculosis. *Clin Infect Dis* 2022;**25**:25.	Outcome	Software eligible (Qxr) but not AI+clinician, no nodule or lung cancer outcomes

http://dx.doi.org/10.1093/cid/ciac679		
Kapoor A, Kapoor A, Mahajan G. Use of Artificial Intelligence to Triage Patients with Flu-Like Symptoms Using Imaging in Non-COVID-19 Hospitals during COVID-19 Pandemic: An Ongoing 8-Month Experience. *Indian J Radiol Imaging* 2021;**31**(4):901-9. http://dx.doi.org/10.1055/s-0041-1741103	Outcome	Not a named software (COVID-19 AI, Quibim) or with clinician
Kim S, Lin CW, Tseng GC. MetaKTSP: a meta-analytic top scoring pair method for robust cross-study validation of omics prediction analysis. *Bioinformatics* 2016;**32**(13):1966-73. http://dx.doi.org/10.1093/bioinformatics/btw115	Outcome	Not a named software, not AI for CXR, no radiologists, no relevant outcomes
Kim W, Lee SM, Kim JI, Ahn Y, Park S, Choe J*, et al.* Utility of a Deep Learning Algorithm for Detection of Reticular Opacity on Chest Radiography in Patients With Interstitial Lung Disease. *AJR American Journal of Roentgenology* 2022;**218**(4):642-50. http://dx.doi.org/10.2214/AJR.21.26682	Outcome	Software eligible (VUNO Med-Chest X-Ray, VUNO) and AI+clinican vs clinican, no outcomes of nodules or lung cancer, those with nodules were excluded from the CXRs being investigated
Kör H, Erbay H, Yurttakal AH. Diagnosing and differentiating viral pneumonia and COVID-19 using X-ray images. *Multimedia Tools Appl* 2022;**81**(27):39041–57. http://dx.doi.org/10.1007/s11042-022-13071-z	Outcome	Not eligible software, pneumonia and covid, no cancer outcomes
Lee YW, Huang SK, Chang RF. CheXGAT: A disease correlation-aware network for thorax disease diagnosis from chest X-ray images. *Artif Intell Med* 2022;**132 (no pagination)**. http://dx.doi.org/10.1016/j.artmed.2022.102382	Intervention	Not eligible software, population unclear, from CHestX-ray8 dataset; thorax disease
Louati H, Louati A, Bechikh S, Masmoudi F, Aldaej A, Kariri E. Topology optimization search of deep convolution neural networks for CT and X-ray image classification. *BMC Med Imaging* 2022;**22**(1):120. http://dx.doi.org/10.1186/s12880-022-00847-w	Outcome	Not a named commercial software, not AI+clinician, single arm study compared with historical studies, no relevant outcomes
MacPherson M, Muthuswamy K, Amlani A, Hutchinson C, Goh V, Montana G. Assessing the Performance of Automated Prediction and Ranking of Patient Age from Chest X-rays Against Clinicians. Medical Image Computing and Computer Assisted Intervention – MICCAI 2022; Cham, abstract no. 302, p. 255-65.	Intervention	Not a symptomatic or asymptomatic population looking for lung cancer / nodules, not commercial software, not AI+clinician vs clinician, no outcomes
Maheshwari S, Sharma RR, Kumar M. LBP-based information assisted intelligent system for COVID-19 identification. *Comput Biol Med* 2021;**134**(C):8. http://dx.doi.org/10.1016/j.compbiomed.2021.104453	Outcome	Not eligible software, covid, no cancer outcomes
Manokaran J, Zabihollahy F, Hamilton-Wright A, Ukwatta E. Detection of COVID-19 from chest x-ray images using transfer learning. *Journal of Medical Imaging* 2021;**8(S1) (no pagination)**. http://dx.doi.org/10.1117/1.JMI.8.S1.017503	Outcome	Not eligible software, covid, no cancer outcomes
Masud M. A light-weight convolutional Neural Network Architecture for classification of COVID-19 chest X-Ray images. *Multimedia Systems* 2022;**28**(4):1165-74. http://dx.doi.org/10.1007/s00530-021-00857-8	Outcome	Not an eligible software or outcomes
Mushtaq J, Pennella R, Lavalle S, Colarieti A, Steidler S, Martinenghi CMA*, et al.*	Intervention	Software eligible (Qxr) but AI alone. Participants

Initial chest radiographs and artificial intelligence (AI) predict clinical outcomes in COVID-19 patients: analysis of 697 Italian patients. *Eur Radiol* 2021;**31**(3):1770-9. http://dx.doi.org/10.1007/s00330-020-07269-8		presenting to ED with positive Covid-19 test, AI vs clinician, no relevant outcomes
Nash M, Kadavigere R, Andrade J, Sukumar CA, Chawla K, Shenoy VP*, et al.* Deep learning, computer-aided radiography reading for tuberculosis: a diagnostic accuracy study from a tertiary hospital in India. *Sci Rep* 2020;**10**(1):210. http://dx.doi.org/10.1038/s41598-019-56589-3	Intervention	Unclear if referred from primary care, software eligible (qXR) is stand alone, outcomes not nodules (but included in 'opacity') or cancer
Patel BN, Rosenberg L, Willcox G, Baltaxe D, Lyons M, Irvin J*, et al.* Human-machine partnership with artificial intelligence for chest radiograph diagnosis. *NPJ Digit Med* 2019;**2**:111. http://dx.doi.org/10.1038/s41746-019-0189-7	Intervention	Not relevant software.
Qin ZZ, Ahmed S, Sarker MS, Paul K, Adel ASS, Naheyan T*, et al.* Tuberculosis detection from chest x-rays for triaging in a high tuberculosis-burden setting: an evaluation of five artificial intelligence algorithms. *The Lancet Digital Health* 2021;**3**(9):e543-e54. http://dx.doi.org/10.1016/S2589-7500(21)00116-3	Outcome	Eligible software (InferRead DR, Lunit INSIGHT, qXR and 2 others not eligible) but stand alone, vs radiologist; population presented or referred for tuberculosis screening but unclear if primary care referrals; no nodule or cancer outcomes
Qin ZZ, Sander MS, Rai B, Titahong CN, Sudrungrot S, Laah SN*, et al.* Using artificial intelligence to read chest radiographs for tuberculosis detection: A multi-site evaluation of the diagnostic accuracy of three deep learning systems. *Sci Rep* 2019;**9**(1):15000. http://dx.doi.org/10.1038/s41598-019-51503-3	Outcome	Eligible software (Lunit INSIGHT, qXR and one not relevant) but appears to be AI alone. Referral status unclear but enrolled in outpatient dept. No relevant outcomes
Rao PS, Bheemavarapu P, Kalyampudi PSL, Rao TVM. An Efficient Method for Coronavirus Detection Through X-rays Using Deep Neural Network. *Current Medical Imaging* 2022;**18(6)**:587-92. http://dx.doi.org/10.2174/1573405617999210112193220	Outcome	Not eligible software, covid, no cancer outcomes
Rathi R, Balayan N, Kumar CNSV. Pneumonia detection using chest X-ray. *International Journal of Pharmaceutical Research* 2020;**12(3)**:1150-3. http://dx.doi.org/10.31838/ijpr/2020.12.03.181	Outcome	Not eligible software, covid, no cancer outcomes
Reis HC, Turk V. COVID-DSNet: A novel deep convolutional neural network for detection of coronavirus (SARS-CoV-2) cases from CT and Chest X-Ray images. *Artif Intell Med* 2022;**134 (no pagination)**. http://dx.doi.org/10.1016/j.artmed.2022.102427	Outcome	Not eligible software, covid, no cancer outcomes
Salama WM, Shokry A, Aly MH. A generalized framework for lung Cancer classification based on deep generative models. *Multimedia Tools Appl* 2022;**81**(23):32705–22. http://dx.doi.org/10.1007/s11042-022-13005-9	Intervention	Not eligible software, population unclear, cancer detection
Santosh KC, Antani S. Automated Chest X-Ray Screening: Can Lung Region Symmetry Help Detect Pulmonary Abnormalities? *IEEE Trans Med Imaging* 2018;**37**(5):1168-77. http://dx.doi.org/10.1109/TMI.2017.2775636	Outcome	Not AI and no outcomes

Singh A, Lall B, Panigrahi BK, Agrawal A, Thangakunam B, Christopher DJ. Deep LF-Net: Semantic lung segmentation from Indian chest radiographs including severely unhealthy images. *Biomed Signal Process Control* 2021;**68 (no pagination)**. http://dx.doi.org/10.1016/j.bspc.2021.102666	Outcome	Not eligible software, databases of different diseases including nodules, no relevant outcomes
Sun W, Wu D, Luo Y, Liu L, Zhang H, Wu S*, et al.* A Fully Deep Learning Paradigm for Pneumoconiosis Staging on Chest Radiographs. *IEEE Journal of Biomedical and Health Informatics* 2022;**26(10)**:5154-64. http://dx.doi.org/10.1109/JBHI.2022.3190923	Outcome	Not eligible software, pneumoconiosis, no cancer outcomes
Sung J, Park S, Lee SM, Bae W, Park B, Jung E*, et al.* Added value of deep learning-based detection system for multiple major findings on chest radiographs: A randomized crossover study. *Radiology* 2021;**299(2)**:450-9. http://dx.doi.org/10.1148/RADIOL.2021202818	Population	Eligible software Med-Chest X-Ray (Vuno) + radiologist; population is inpatients + outpatients, proportions not reported, referral and symptom status unclear, proportion unclear; limited relevant outcomes but includes nodule detection
Tan M, Deklerck R, Cornelis J, Jansen B. Phased searching with NEAT in a Time-Scaled Framework: Experiments on a computer-aided detection system for lung nodules. *Artif Intell Med* 2013;**59**(3):157–67. http://dx.doi.org/10.1016/j.artmed.2013.07.002	Intervention	Not eligible software, population unclear, nodule detection
Tavaziva G, Majidulla A, Nazish A, Saeed S, Benedetti A, Khan AJ*, et al.* Diagnostic accuracy of a commercially available, deep learning-based chest X-ray interpretation software for detecting culture-confirmed pulmonary tuberculosis. *Int J Infect Dis* 2022;**122**:15-20. http://dx.doi.org/10.1016/j.ijid.2022.05.037	Outcome	Eligible software (Lunit Insight) but appears to be stand alone; people presenting with tuberculosis symptoms or household contacts, unclear if referred from primary care; no nodule or cancer outcomes
Vajda S, Karargyris A, Jaeger S, Santosh KC, Candemir S, Xue Z*, et al.* Feature Selection for Automatic Tuberculosis Screening in Frontal Chest Radiographs. *J Med Syst* 2018;**42**(8):146. http://dx.doi.org/10.1007/s10916-018-0991-9	Outcome	Not AI, no outcomes
Vieira P, Sousa O, Magalhães D, Rabêlo R, Silva R. Detecting pulmonary diseases using deep features in X-ray images. *Pattern Recogn* 2021;**119**(C):13. http://dx.doi.org/10.1016/j.patcog.2021.108081	Outcome	Not eligible software, covid and pneumonia, no relevant outcomes
Wang K, Zhang X, Huang S, Chen F. Automatic Detection of Pneumonia in Chest X-Ray Images Using Cooperative Convolutional Neural Networks. Paper presented at: Pattern Recognition and Computer Vision: Second Chinese Conference, PRCV 2019, Xi'an, China, November 8–11, 2019, Proceedings, Part II; Xi'an, China. URL: https://doi.org/10.1007/978-3-030-31723-2_28	Outcome	Not eligible software, pneumonia, no relevant outcomes
Zaglam N, Jouvet P, Flechelles O, Emeriaud G, Cheriet F. Computer-aided diagnosis system for the Acute Respiratory Distress Syndrome from chest radiographs. *Comput Biol Med* 2014;**52**:41–8. http://dx.doi.org/10.1016/j.compbiomed.2014.06.006	Outcome	Not eligible software, acute respiratory distress syndrome, no relevant outcomes

Zhang R, Tie X, Qi Z, Bevins NB, Zhang C, Griner D, *et al.* Diagnosis of Coronavirus Disease 2019 Pneumonia by Using Chest Radiography: Value of Artificial Intelligence. *Radiology* 2021;**298(2)**:E88-E97. http://dx.doi.org/10.1148/RADIOL.2020202944	Outcome	Not eligible software, covid, no relevant outocmes
Zhou W, Cheng G, Zhang Z, Zhu L, Jaeger S, Lure FYM, et al. Deep learning-based pulmonary tuberculosis automated detection on chest radiography: large-scale independent testing. Quantitative Imaging in Medicine & Surgery 2022;12(4):2344-55. http://dx.doi.org/10.21037/qims-21-676	Intervention	Not an eligible software, symptom and referal status unclear, focus is on TB

Ongoing studies

Title and link	Reason for exclusion	Comments
Retrospective Study of Carebot AI CXR Performance in Preclinical Practice	Intervention	Not commercial named software, population in hospital, not AI+radiologist
Research and development of artificial intelligence assistant diagnosis and clinical decision system for pulmonary ground glass nodules	Intervention	Not commercial named software, not Xray, population pre surgery or biopsy
Diagnosis of lung nodule and lung cancer on screening chest radiographs: Comparative clinical trial for evaluation of artificial intelligence-integrated PACS versus conventional PACS	Duplicate	In database searches, already screened.
Sensitivity of chest X-ray in patients with suspected acute thoracic diseases in emergency department: Randomized controlled trial to assess efficiency of artificial intelligence-based computer-aided detection system	Duplicate	In database searches, already screened.
To compare the outcome performance of Digital Chest Radiograph and radiologist diagnosis based on chest x ray	Intervention	Exclude – software not named, not AI+radiologist (AI vs radiologist reference standard)
A Study to Assess the Impact of an Artificial Intelligence (AI) System on Chest X-ray Reporting	Duplicate	Found via other sources, already screened.
A study to evaluate the effectiveness of computer artificial inteligence in identifying and classifying abnormalites in chest radiographs	Intervention	Software unclear but no manufacturer specified, non-commercial funding. Includes AI+radiologist vs radiologist alone. Population could include lung cancer but suggests will be subgroup analyses. Study not yet recruiting
Clinical Validation of Machine Learning Triage of Chest Radiographs	Intervention	Software unclear but no manufacturer specified + non-commercial funding.
Artificial Intelligence to Improve Physicians' Interpretation of Chest X-Rays in Breathless Patients	Intervention	Not commercial named software, population presenting to A+E

Multicenter Validation Study of an Artificial Intelligence Tool for Automatic Classification of Chest X-rays	Intervention	Not commercial named software, not nodules or lung cancer outcomes
Use of artificial intelligence to interpret chest X-rays	Intervention	Named commercial software (Qure.ai), unclear population and not AI+radiologist, validation study
Evaluation of Computer-Assisted-Detection (CAD) software for Chest X-ray lung Nodule	Intervention	Not a named software (sponsor is lpixel inc), unclear if AI+radiologist. Study is completed and URL links to study already screened
Deep Learning Model for Pure Solid Nodules Classification	Intervention	CT
Deep Learning Signature for Predicting Occult Nodal Metastasis of Clinical N0 Lung Cancer	Intervention	PET/CT
Effects of percutaneous vertebroplasty on respiratory parameters in patients with osteoporotic vertebral compression fractures	Intervention	Percutaneous vertebroplasty
Research on Differential Diagnosis of Pulmonary Nodules Based on Radiomics and Artificial Intelligence	Intervention	CT
Development and validation of AI model to predict the surgical site infection in lung cancer surgery	Intervention	AI-based model to predict the outcome of surgical site infection (SSI)
Constructing a deep learning model for the differentiation of benign and malignant single solid small nodules based on multi-omics features: a prospective, multi-center clinical study	Intervention	CT
Future Health Today: A cluster randomised controlled trial of quality improvement activities in general practice	Intervention	Quality improvement activities in general practice. Technology platform consisting of audit, point of care clinical decision support, and QI templates. Education activities.
A Preliminary Study on the Detection of Plasma Markers in Early Diagnosis for Lung Cancer	Intervention	Machine-learning for Detection of Plasma Markers
Research on the rapid pathological diagnosis of lung nodules based on intraoperative tumor images and preoperative CT images based on deep learning	Intervention	CT
Establishment and Application Research of Early Lung Cancer Prognosis Grading System Based on Machine Learning and New Pathological Features	Intervention	Grading System Based on Machine Learning and New Pathological Features
Automatic PD-L1 immunohistochemistry evaluation system for non-small cell lung cancer based on deep learning	Intervention	immunohistochemistry
Application Research of Artificial Intelligence Assistant System in in Comprehensive Preoperative Evaluation of Early Lung Cancer	Intervention	AI in Preoperative Evaluation. CT already used
The application of artificial intelligence diagnosis system in the pathological greades differentiation in lung adenocarcinoma	Intervention	Deep learning in the pathological grades differentiation
Multicenter clinical study of PET / CT radiomics in predicting gene mutation of lung cancer	Intervention	PET/CT

Study for the combination of medical image artificial intelligence technology and intraoperative frozen pathology can accurately predict the stage and subtypes of lung adenocarcinoma	Intervention	CT and pathological diagnosis
Tumor Invasiveness Estimation of Artificial Intelligence System for Subsolid Nodules on Computed Tomography: Diagnostic Performance and Utility Verification in Clinical Practice	Intervention	CT
Technical standard and application of intelligence assisted ultrasound indiagnosis of subpleural lung lesions	Intervention	ultrasound
Lung Nodule Imaging Biobank for Radiomics and AI Research LIBRA	Intervention	CT
Investigation of the BRAF mutation status in the pleural punctate in patients with malignant melanoma, colorectal or lung cancer, that show a BRAF mutation in the primary tumor and the comparison of the result with the conventional cytological / immunohistochemical and molecular cytology findings of the pleural punctate	Intervention	Automatic diagnostic system for malignant pleural effusion
Best Start Trial: early intervention physiotherapy to improve motor outcomes in infants at high risk of cerebral palsy or motor delay	Intervention	Physiotherapy
China Lung Cancer Screening (CLUS) Study Version 2.0	Intervention	CT
Development and Validation of a Three-dimensional Convolutional Neural Network for Automated Detection of Lung Nodule from Computed Tomography Images	Intervention	CT
Evaluation of Use of Diagnostic AI for Lung Cancer in Practice	Intervention	CT (specified in clinicaltrials.gov record)
Diagnostic Performance of Neural Network-Based Artificial Intelligent in Detecting Pulmonary Nodule on Chest CT	Intervention	CT
CT data collection of pulmonary nodules and analysis of artificial intelligence	Intervention	CT
Follow up after curative-intent lung cancer treatment	Intervention	PET-CT or CT surveillance of lung cancer patients
A Phase III Study of MEDI4736, given as Monotherapy or in Combination with Tremelimumab, versus Standard of Care in Patients with Locally Advanced or Metastatic Non-Small Cell Lung Cancer	Intervention	pharmaceutical
Screening for Early Lung Cancer in Shanghai, China	Intervention	CT
Natural History of Lung Nodules Seen on CT Scans From Participants at High-Risk of Developing Lung Cancer	Intervention	CT
Development and Validation of Artificial Intelligence Based Tool to read Chest X-rays in order to detect Pulmonary TB and other lung diseases	Intervention	Software unclear but no manufacturer specified, non-commercial funding, likely screening population, tuberculosis
Feasibility of AI-based Heart Function Prediction Model Using CXR	Population	Patient who visited the emergency room or outpatient clinic due to dyspnea and chest Pain. Outcome Left Ventricular Ejection Fraction < 40%
Rapid Research in Diagnostics Development for TB Network	Intervention	Software not AI, no nodules as outcomes, tuberculosis

Comparative Study of Artificial Intelligence and Radiologists in Assessing Severity of COVID19 Patient Images	Intervention	CT
Novel Artificial Intelligence Algorithm to screen COVID-19 Patients from X-Ray , CT-Scan of Thorax and Voice Sampling through Android App and storage through Cloud	Intervention	Software unclear but no manufacturer specified, non-commercial funding
Development and evaluation of minimally invasive and Dynamic digital radiography	Intervention	Not AI. CXR for evaluation of pulmonary function
Artificial Intelligence Algorithms for Discriminating Between COVID-19 and Influenza Pneumonitis Using Chest X-Rays	Intervention	Software unclear but not a named manufacturer, no outcomes of relevance
Study for the key issues of the diagnosis and treatment of novel coronavirus pneumonia (COVID-19) based on the medical imaging	Intervention	Not AI algorithm / software, no outcomes
Can periodically promoting tuberculosis and HIV testing reduce undiagnosed infectious tuberculosis and tuberculosis transmission in communities?	Intervention	Not AI, no outcomes, population likely screening
MAchine Learning in whole Body Oncology	Intervention	Magnetic resonance (MR) imaging
Clinical trial to evaluate the clinical effectiveness of pneumothorax reading results using artificial intelligence software in chest X-ray images	Intervention	Software unclear but no manufacturer specified, non-commercial funding, no outcomes of relevance
Accuracy of artificial intelligence in CXR screening for pulmonary tuberculosis	Intervention	Not commercial named software
Crowdsourcing an Open COVID-19 Imaging Repository for AI Research	Intervention	Not commercial named software
Predicting morbidity and mortality of preterm infants by analyzing chest x-ray images at admission using deep learning algorithms	Population	Preterm infants
Evaluation of detection of Pulmonary TB by computer aided technology	Intervention	Not named commercial software but Qure.ai are co-sponsors, population referral unclear, not AI+radiologist, no outcomes
The retrospective study for the development of the Artificial Intelligence (AI) regarding with the chest x-ray in pulmonary arterial hypertension	Intervention	Not commercial named software (3M)
Potential of Deep Learning in Assessing Pneumoconiosis Depicted on Digital Chest Radiography	Intervention	Not commercial named software, non-commercial funding, screening population
Evaluation of Pneumoconiosis High Risk Early Warning Models	Intervention	Not commercial named software
Coronavirus: Ventilator Outcomes Using Artificial Intelligence Chest Radiographs & Other Evidence-based Co-variates	Intervention	Not commercial named software, population in hospital
Evaluation of a COVID-19 Pneumonia CXR AI Detection Algorithm	Intervention	Not commercial named software
Advantage of Artificial Intelligence to detect COVID 19 using Chest X-Ray.	Intervention	Not commercial named software
Software for COVID19 Detection from Chest X-Ray, CT or Ultrasonography	Intervention	Not commercial named software
Classification of COVID-19 Infection in Posteroanterior Chest X-rays	Intervention	Not commercial named software

Reviews used for reference checking

Aggarwal R, Sounderajah V, Martin G, Ting DSW, Karthikesalingam A, King D, et al. Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. Npj Digital Medicine 2021;4(1):65. http://dx.doi.org/10.1038/s41746-021-00438-z
Amir GJ, Lehmann HP. After Detection: The Improved Accuracy of Lung Cancer Assessment Using Radiologic Computer-aided Diagnosis. Acad Radiol 2016;23(2):186-91. http://dx.doi.org/10.1016/j.acra.2015.10.014
Forte GC, Altmayer S, Silva RF, Stefani MT, Libermann LL, Cavion CC, et al. Deep Learning Algorithms for Diagnosis of Lung Cancer: A Systematic Review and Meta-Analysis. Cancers (Basel) 2022;14(16):09. http://dx.doi.org/10.3390/cancers14163856
Haber M, Drake A, Nightingale J. Is there an advantage to using computer aided detection for the early detection of pulmonary nodules within chest X-Ray imaging? Radiography (London) 2020;26(3):e170-e8. http://dx.doi.org/10.1016/j.radi.2020.01.002
Lee JH, Hwang EJ, Kim H, Park CM. A narrative review of deep learning applications in lung cancer research: from screening to prognostication. Translational Lung Cancer Research 2022;11(6):1217-29. http://dx.doi.org/10.21037/tlcr-21-1012
Li D, Pehrson LM, Lauridsen CA, Tottrup L, Fraccaro M, Elliott D, et al. The Added Effect of Artificial Intelligence on Physicians' Performance in Detecting Thoracic Pathologies on CT and Chest X-ray: A Systematic Review. Diagnostics 2021;11(12):26. http://dx.doi.org/10.3390/diagnostics11122206
Qin C, Yao D, Shi Y, Song Z. Computer-aided detection in chest radiography based on artificial intelligence: a survey. Biomed Eng Online 2018;17(1):113. http://dx.doi.org/10.1186/s12938-018-0544-y

Reviews not selected for reference checking

Cancer reviews
Jones CM, Buchlak QD, Oakden-Rayner L, Milne M, Seah J, Esmaili N, et al. Chest radiographs and machine learning - Past, present and future. J Med Imaging Radiat Oncol 2021;65(5):538-44. http://dx.doi.org/10.1111/1754-9485.13274
Ley S, Ley-Zaporozhan J. Novelties in imaging in pulmonary fibrosis and nodules. A narrative review. Pulmonology 2020;26(1):39-44. http://dx.doi.org/10.1016/j.pulmoe.2019.09.009
Meedeniya D, Kumarasinghe H, Kolonne S, Fernando C, Diez IT, Marques G. Chest X-ray analysis empowered with deep learning: A systematic review. Applied Soft Computing 2022;126:109319. http://dx.doi.org/10.1016/j.asoc.2022.109319
Mercy Theresa M, Bharathi VS. A Survey on CAD technique for various abnormality classification in chest radiography. Research Journal of Pharmaceutical, Biological and Chemical Sciences 2016;7(4):331-42. http://www.rjpbcs.com/
Ozcelik N, Selimoglu I. Artificial intelligence applications in pulmonology and its advantages during the pandemic period. Tuberkuloz ve Toraks 2021;69(3):380-6. http://dx.doi.org/10.5578/tt.20219710
Subramanian N, Elharrouss O, Al-Maadeed S, Chowdhury M. A review of deep learning-based detection methods for COVID-19. Comput Biol Med 2022;143:105233. http://dx.doi.org/10.1016/j.compbiomed.2022.105233
Suzuki K. Overview of deep learning in medical imaging. Radiological Physics and Technology 2017;10(3):257-73. http://dx.doi.org/10.1007/s12194-017-0406-5
Tufail AB, Ma YK, Kaabar MKA, Martinez F, Junejo AR, Ullah I, et al. Deep Learning in Cancer Diagnosis and Prognosis Prediction: A Minireview on Challenges, Recent Trends, and Future Directions. Comput Math Methods Med 2021;2021 (no pagination). http://dx.doi.org/10.1155/2021/9025470
Yang J, Wang H, Geng C, Dai Y, Ji J. Advances in intelligent diagnosis methods for pulmonary ground-glass opacity nodules. Biomed Eng Online 2018;17(1) (no pagination). http://dx.doi.org/10.1186/s12938-018-0435-2

Yang Y, Feng X, Chi W, Li Z, Duan W, Liu H, et al. Deep learning aided decision support for pulmonary nodules diagnosing: a review. J Thorac Dis 2018;10(Suppl 7):S867-S75. http://dx.doi.org/10.21037/jtd.2018.02.57
Zheng X, He B, Hu Y, Ren M, Chen Z, Zhang Z, et al. Diagnostic Accuracy of Deep Learning and Radiomics in Lung Cancer Staging: A Systematic Review and Meta-Analysis. Frontiers in Public Health 2022;10:938113. http://dx.doi.org/10.3389/fpubh.2022.938113
Non-cancer reviews
Adamidi ES, Mitsis K, Nikita KS. Artificial intelligence in clinical care amidst COVID-19 pandemic: A systematic review. *Computational and Structural Biotechnology Journal* 2021;**19**:2833-50. http://dx.doi.org/10.1016/j.csbj.2021.05.010
Alzubaidi M, Zubaydi HD, Bin-Salem AA, Abd-Alrazaq AA, Ahmed A, Househ M. Role of deep learning in early detection of COVID-19: Scoping review. *Computer Methods and Programs in Biomedicine Update* 2021;**1**:100025. http://dx.doi.org/10.1016/j.cmpbup.2021.100025
Chan J, Auffermann WF. Artificial Intelligence in the Imaging of Diffuse Lung Disease. *Radiol Clin North Am* 2022;**60(6)**:1033-40. http://dx.doi.org/10.1016/j.rcl.2022.06.014
Corbacho Abelaira MD, Corbacho Abelaira F, Ruano-Ravina A, Fernandez-Villar A. Use of Conventional Chest Imaging and Artificial Intelligence in COVID-19 Infection. A Review of the Literature. *Open Respiratory Archives* 2021;**3(1) (no pagination)**. http://dx.doi.org/10.1016/j.opresp.2020.100078
Fusco R, Grassi R, Granata V, Setola SV, Grassi F, Cozzi D*, et al.* Artificial Intelligence and COVID-19 Using Chest CT Scan and Chest X-ray Images: Machine Learning and Deep Learning Approaches for Diagnosis and Treatment. *Journal of Personalized Medicine* 2021;**11**(10):30. http://dx.doi.org/10.3390/jpm11100993
Ge C, Zhang L, Xie L, Kong R, Zhang H, Chang S. COVID-19 Imaging-based AI Research-A Literature Review. *Current Medical Imaging* 2022;**18(5)**:496-508. http://dx.doi.org/10.2174/1573405617666210902103729
Ghaderzadeh M, Aria M, Asadi F. X-Ray Equipped with Artificial Intelligence: Changing the COVID-19 Diagnostic Paradigm during the Pandemic. *BioMed Research International* 2021;**2021**:9942873. http://dx.doi.org/10.1155/2021/9942873
Ghaderzadeh M, Asadi F. Deep Learning in the Detection and Diagnosis of COVID-19 Using Radiology Modalities: A Systematic Review. *J Healthc Eng* 2021;**2021 (no pagination)**. http://dx.doi.org/10.1155/2021/6677314
Ghrabli S, Elgendi M, Menon C. Challenges and Opportunities of Deep Learning for Cough-Based COVID-19 Diagnosis: A Scoping Review. *Diagnostics* 2022;**12(9) (no pagination)**. http://dx.doi.org/10.3390/diagnostics12092142
Gillman AG, Lunardo F, Prinable J, Belous G, Nicolson A, Min H*, et al.* Automated COVID-19 diagnosis and prognosis with medical imaging and who is publishing: a systematic review. *Physical and Engineering Sciences in Medicine* 2022;**45**(1):13-29. http://dx.doi.org/10.1007/s13246-021-01093-0
Harris M, Qi A, Jeagal L, Torabi N, Menzies D, Korobitsyn A*, et al.* A systematic review of the diagnostic accuracy of artificial intelligence-based computer programs to analyze chest x-rays for pulmonary tuberculosis. *PLoS ONE [Electronic Resource]* 2019;**14**(9):e0221339. http://dx.doi.org/10.1371/journal.pone.0221339
Ito R, Iwano S, Naganawa S. A review on the use of artificial intelligence for medical imaging of the lungs of patients with coronavirus disease 2019. *Diagn Interv Radiol* 2020;**26**(5):443-8. http://dx.doi.org/10.5152/dir.2019.20294
Jia LL, Zhao JX, Pan NN, Shi LY, Zhao LP, Tian JH*, et al.* Artificial intelligence model on chest imaging to diagnose COVID-19 and other pneumonias: A systematic review and meta-analysis. *European Journal of Radiology Open* 2022;**9**:100438. http://dx.doi.org/10.1016/j.ejro.2022.100438
Karthik R, Menaka R, Hariharan M, Kathiresan GS. AI for COVID-19 Detection from Radiographs: Incisive Analysis of State of the Art Techniques, Key Challenges and Future Directions. *Irbm* 2022;**43(5)**:486-510. http://dx.doi.org/10.1016/j.irbm.2021.07.002
Khan FA, Pande T, Tessema B, Song R, Benedetti A, Pai M*, et al.* Computer-aided reading of tuberculosis chest radiography: Moving the research agenda

forward to inform policy. *Eur Respir J* 2017;**50(1) (no pagination)**. http://dx.doi.org/10.1183/13993003.00953-2017
Komolafe TE, Cao Y, Nguchu BA, Monkam P, Olaniyi EO, Sun H*, et al.* Diagnostic Test Accuracy of Deep Learning Detection of COVID-19: A Systematic Review and Meta-Analysis. *Acad Radiol* 2021;**28(11)**:1507-23. http://dx.doi.org/10.1016/j.acra.2021.08.008
Kriza C, Amenta V, Zenie A, Panidis D, Chassaigne H, Urban P*, et al.* Artificial intelligence for imaging-based COVID-19 detection: Systematic review comparing added value of AI versus human readers. *Eur J Radiol* 2021;**145**:110028. http://dx.doi.org/10.1016/j.ejrad.2021.110028
Kufel J, Bargiel K, Kozlik M, Czogalik L, Dudek P, Jaworski A*, et al.* Application of artificial intelligence in diagnosing COVID-19 disease symptoms on chest X-rays: A systematic review. *Int J Med Sci* 2022;**19**(12):1743-52. http://dx.doi.org/10.7150/ijms.76515
Li Y, Zhang Z, Dai C, Dong Q, Badrigilan S. Accuracy of deep learning for automated detection of pneumonia using chest X-Ray images: A systematic review and meta-analysis. *Comput Biol Med* 2020;**123**:103898. http://dx.doi.org/10.1016/j.compbiomed.2020.103898
Lopez-Cabrera JD, Orozco-Morales R, Portal-Diaz JA, Lovelle-Enriquez O, Perez-Diaz M. Current limitations to identify covid-19 using artificial intelligence with chest x-ray imaging (part ii). The shortcut learning problem. *Health & Technology* 2021;**11**(6):1331-45. http://dx.doi.org/10.1007/s12553-021-00609-8
Montazeri M, ZahediNasab R, Farahani A, Mohseni H, Ghasemian F. Machine Learning Models for Image-Based Diagnosis and Prognosis of COVID-19: Systematic Review. *JMIR Medical Informatics* 2021;**9**(4):e25181. http://dx.doi.org/10.2196/25181
Negrini D, Danese E, Henry BM, Lippi G, Montagnana M. Artificial intelligence at the time of COVID-19: who does the lion's share? *Clin Chem Lab Med* 2022;**60**(12):1881-6. http://dx.doi.org/10.1515/cclm-2022-0306
Oloko-Oba M, Viriri S. A Systematic Review of Deep Learning Techniques for Tuberculosis Detection From Chest Radiograph. *Frontiers in Medicine* 2022;**9**:830515. http://dx.doi.org/10.3389/fmed.2022.830515
Rahmani AM, Azhir E, Naserbakht M, Mohammadi M, Aldalwie AHM, Majeed MK*, et al.* Automatic COVID-19 detection mechanisms and approaches from medical images: a systematic review. *Multimedia Tools & Applications* 2022;**81**(20):28779-98. http://dx.doi.org/10.1007/s11042-022-12952-7
Rezayi S, Ghazisaeedi M, Kalhori SRN, Saeedi S. Artificial Intelligence Approaches on X-ray-oriented Images Process for Early Detection of COVID-19. *Journal of Medical Signals & Sensors* 2022;**12**(3):233-53. http://dx.doi.org/10.4103/jmss.jmss_111_21
Santosh KC, Allu S, Rajaraman S, Antani S. Advances in Deep Learning for Tuberculosis Screening using Chest X-rays: The Last 5 Years Review. *J Med Syst* 2022;**46**(11):82. http://dx.doi.org/10.1007/s10916-022-01870-8
Sinwar D, Dhaka VS, Tesfaye BA, Raghuwanshi G, Kumar A, Maakar SK*, et al.* Artificial Intelligence and Deep Learning Assisted Rapid Diagnosis of COVID-19 from Chest Radiographical Images: A Survey. *Contrast Media Mol Imaging* 2022;**2022**:1306664. http://dx.doi.org/10.1155/2022/1306664
Tavaziva G, Harris M, Abidi SK, Geric C, Breuninger M, Dheda K*, et al.* Chest X-ray Analysis With Deep Learning-Based Software as a Triage Test for Pulmonary Tuberculosis: An Individual Patient Data Meta-Analysis of Diagnostic Accuracy. *Clin Infect Dis* 2022;**74**(8):1390-400. http://dx.doi.org/10.1093/cid/ciab639
Ting DSW, Tan TE, Lim CCT. Development and Validation of a Deep Learning System for Detection of Active Pulmonary Tuberculosis on Chest Radiographs: Clinical and Technical Considerations. *Clin Infect Dis* 2019;**69(5)**:748-50. http://dx.doi.org/10.1093/cid/ciy969
Wang L, Zhang Y, Wang D, Tong X, Liu T, Zhang S*, et al.* Artificial Intelligence for COVID-19: A Systematic Review. *Frontiers in Medicine* 2021;**8 (no pagination)**. http://dx.doi.org/10.3389/fmed.2021.704256