CONFIDENTIAL UNTIL PUBLISHED

Assessment Group's Report

EarlyCDT Lung for lung cancer risk classification of solid pulmonary nodules

Produced by CRD and CHE Technology Assessment Group, University of York, Heslington, York, [redacted]

Authors Ana Duarte, Research Fellow, CHE

Mark Corbett, Research Fellow, CRD

Hollie Melton, Research Fellow, CRD

Melissa Harden, Information Specialist, CRD

Stephen Palmer, Professor, CHE

Marta Soares, Senior Research Fellow, CHE

Mark Simmonds, Senior Research Fellow, CRD

Correspondence to Mark Simmonds

[redacted]

[redacted]

Date completed 16 August 2021

Source of funding

This report was commissioned by the NIHR Evidence Synthesis Programme as project number DAP54 (NIHR13/34/96).

Declared competing interests of the authors

The authors have no competing interests to declare

Acknowledgements

Prof. Matthew Callister (Specialist committee member, Consultant in Respiratory Medicine, Leeds Teaching Hospital NHS Trust) provided expert clinical advice to the EAG throughout the project.

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:

Duarte A, Corbett M, Melton H, Harden M, Palmer S, Soares M, Simmonds M. EarlyCDT Lung for lung cancer risk classification of solid pulmonary nodules: A Diagnostics Assessment Report. York EAG, 2021

Contributions of authors

Ana Duarte: Performed cost-effectiveness reviews and wrote economic sections of this report

Mark Corbett: Performed the systematic review and wrote associated sections of the report

Hollie Melton: Performed the systematic review and contributed to associated sections of the report

Melissa Harden: Performed all database searches

Stephen Palmer: Provided economic advice and commentary for economic sections of the report

Marta Soares: Led the cost-effectiveness components of the project and wrote economic sections of this report

Mark Simmonds: Led the clinical components of the project, performed the statistical analyses and wrote associated sections of the report

Note on the text

This report contains no confidential information of any kind.

Copyright statement

ABSTRACT

Background

EarlyCDT Lung is a blood test that might be of use in assessing the malignancy risk of people with solid pulmonary nodules, who may have lung cancer. The test measures the presence of autoantibodies to a panel of 7 lung cancer associated antigens. Elevated levels of these autoantibodies may indicate malignant disease. Its results might be used to modify the risk of malignancy estimated by existing risk calculators, including the Brock and Herder models.

Objectives

To determine the diagnostic accuracy and the clinical and cost-effectiveness of EarlyCDT Lung. To develop a conceptual model and to identify evidence requirements for a robust cost-effectiveness analysis.

Methods

We performed a systematic review of all evidence on EarlyCDT Lung including all diagnostic accuracy, clinical and cost-effectiveness. We searched MEDLINE and other databases in March 2021. Study quality was assessed with QUADAS-2. Evidence on other components of the pulmonary nodule diagnostic pathway (CT surveillance, Brock risk, Herder risk, PET-CT scans and biopsy) was also reviewed.

Where feasible, meta-analyses of diagnostic accuracy were performed, using random effect meta-analysis and bivariate meta-analyses. Clinical outcomes were synthesised narratively. A simulation study investigated the clinical impact of using EarlyCDT Lung.

Additional reviews of cost-effectiveness studies evaluated i) other diagnostic strategies for lung cancer, and ii) screening approaches for lung cancer. A conceptual model was developed.

Results

We identified 47 clinical publications on EarlyCDT Lung, covering 6 potentially eligible patient cohorts. Five of these cohorts (695 patients) reported diagnostic accuracy data on patients with pulmonary nodules. EarlyCDT Lung on its own was found to have poor diagnostic accuracy with a summary sensitivity of 20.2% (95% CI: 10.5 to 35.5) and specificity of 92.2% (95% CI: 86.2 to 95.8).

No evidence on the clinical impact of EarlyCDT was identified. The simulation study suggested that EarlyCDT Lung might have little benefit when examining lower-risk nodules (e.g. below 10% risk) alongside the Brock risk model. It might potentially have some benefit among intermediate risk nodules (10% to 70% risk) after Herder risk analysis.

Two cost-effectiveness studies on EarlyCDT Lung for pulmonary nodules were identified; none was considered suitable to inform the current decision problem. The additional reviews identified 8 diagnostic cost-effectiveness studies, and 34 screening studies; a sample of 9 screening models were reviewed.

The conceptualisation process identified three core components for a future cost-effectiveness assessment of EarlyCDT Lung: i) the features of the subpopulations and relevant heterogeneity, ii) the way EarlyCDT Lung test results affect subsequent clinical management decisions, and iii) how changes in these decisions can affect outcomes. Evidence to connect these components was sparse, but across all reviewed studies value was established by linking earlier diagnosis to stage progression, and stage shift to final outcomes.

Limitations

The evidence on EarlyCDT Lung in patients with pulmonary nodules was very limited, preventing meta-analyses and economic analyses.

Conclusions

The evidence on EarlyCDT Lung in patients with pulmonary nodules is insufficient to draw any firm conclusions as to its diagnostic accuracy or clinical and economic value. EarlyCDT Lung appears to have poor diagnostic accuracy and so might not improve diagnosis. However, this is uncertain because the evidence is so limited.

SCIENTIFIC SUMMARY

Background

Pulmonary nodules are small growths in the lung, often found when having a chest CT scan. These nodules may be cancerous, and so require treatment. In the UK they are generally managed in accordance with the British Thoracic Society (BTS) Guidelines.

For very small nodules people are discharged with no follow-up. For smaller nodules with lower than 10% risk of malignancy, patients are offered regular surveillance using CT scans. For larger nodules the Brock model is used to assess risk of malignancy. If risk is low (<10%) people will be offered CT surveillance. For higher risk nodules PET-CT is recommended, and the nodule risk is then recalculated using the Herder model. For people with 10-70% risk of malignancy biopsy, excision biopsy or CT surveillance may be used. People with risk over 70% are considered for excision or non-surgical treatment.

EarlyCDT Lung is a blood test manufactured by Oncimmune that could potentially be used to assess the malignancy risk of people at risk of lung cancer. The test measures the presence of 7 autoantibodies. A blood sample is considered to indicate malignancy when at least one of the 7 autoantibodies is elevated above a pre-determined cut-off. Oncimmune proposes that the EarlyCDT Lung test result is used to update a patient's estimated risk of malignancy, with a positive test result increasing the risk.

Objectives

The aim of the project was to appraise existing evidence on the potential clinical and cost-effectiveness of the EarlyCDT Lung test for lung cancer risk classification of solid pulmonary nodules, and to develop a conceptual economic model to provide a common understanding of the evidence requirements and evidence linkages required to undertake a robust cost-effectiveness analysis.

Methods

Diagnostic accuracy and clinical effectiveness

A systematic review was conducted to identify all published studies of EarlyCDT Lung. Comprehensive database searches of MEDLINE, EMBASE and other sources were carried out on 8th March 2021. Further database searching was performed to identify evidence on other parts of the

diagnostic pathway, specifically: Brock and Herder models, CT surveillance, PET-CT scans and biopsy methods.

Key inclusion criteria were:

- Persons with solid pulmonary nodules identified by CT scanning, who may be eligible for further diagnostic testing
- Use of EarlyCDT Lung, or other procedures listed above
- Malignancy confirmed by biopsy or surgical resection; benign nodules confirmed by clinical follow-up of at least one year.
- Studies reported diagnostic accuracy data, or any data on the clinical impact of the technology

Data on study and patient characteristics and results were extracted. Data were also electronically extracted from figures. Data from relevant studies with multiple publications were extracted and reported as a single study. The quality of the diagnostic accuracy studies was assessed using the QUADAS-2 tool.

Given limitations in the evidence a narrative synthesis approach was used, summarising evidence from each study using tables and figures. Meta-analysis of diagnostic accuracy data (sensitivity, specificity and area under the ROC curve) was used where there were sufficient data. Data on diagnostic accuracy for EarlyCDT Lung were combined with data on lung cancer prevalence and nodule risk based on the Brock and Herder models to simulate the potential clinical impact of using EarlyCDT Lung.

Cost-effectiveness

Cost-effectiveness evidence on EarlyCDT Lung for the diagnosis of lung cancer was identified by the abovementioned database searches; the evidence was narratively summarised and tabulated. Studies were appraised for their quality and appropriateness to the decision problem defined by the NICE DAR scope. Additionally, structural and evidential aspects of the decision models were highlighted.

Additional pragmatic literature searches were conducted to identify evidence to support the development of a conceptual model. These searches aimed to identify cost-effectiveness studies evaluating i) other diagnostic strategies for lung cancer, and ii) screening approaches for lung cancer. The studies identified were also narratively summarised to highlight structural and evidential aspects of the decision models (aspects that could be of relevance to the current assessment).

We conceptualised a decision model to inform future evaluations of the cost-effectiveness of use of EarlyCDT Lung, based on the learnings from the literature searches and on clinical advice. The results of the conceptualisation were recorded using influence diagrams, and evidence requirements and uncertainties were highlighted throughout. The conceptualisation process was structured to identify value drivers and value components that could be of relevance for establishing the cost-effectiveness of EarlyCDT Lung in the diagnostic pathway for solid pulmonary nodules.

Results

Systematic review and meta-analysis of EarlyCDT Lung studies

The searches identified a total of 3,233 unique records, of which 47 were included in the review, representing only six distinct patient cohorts. No cohort explicitly performed EarlyCDT Lung after identification of pulmonary nodules. Five of the cohorts reported the diagnostic accuracy of EarlyCDT Lung in patients with nodules. Only two of these five cohorts have been fully published as journal articles. The results from both published cohorts were considered to be at high risk of bias.

The summary sensitivity of EarlyCDT Lung from a bivariate meta-analysis was 20.2% (95% CI 10.5 to 35.5) and the specificity was 92.2% (95% CI 86.2 to 95.8). Based on the summary HSROC curve, Early CDT Lung has around 26% sensitivity at 90% specificity, or 12% sensitivity at 95% specificity. The area under the HSROC curve was 69.4%, suggesting poor to moderate overall diagnostic accuracy. There was little data on diagnostic accuracy by nodule size, or on diagnostic accuracy when combined with other tests, such as Brock risk.

The diagnostic accuracy from the EAG analysis was lower than that claimed by Oncimmune (around 41.3% sensitivity at 90.6% specificity). Consequently, EAG modelling found that the increase in predicted risk of malignancy if Early CDT Lung is positive may be smaller than in the model produced by Oncimmune.

Comparator tests

A meta-analysis of eight studies reporting data on the Brock risk model found it to have very good diagnostic accuracy (AUC 92%, 95% CI: 90% to 95%), but with some evidence of heterogeneity across studies (I^2 = 90%). A meta-analysis of five studies reporting data on the Herder risk model found it to have good diagnostic accuracy overall, with an AUC of 84% (95% CI 77% to 92%). There was substantial heterogeneity (I^2 = 87%).

Although several meta-analyses of the use of PET-CT in patients with pulmonary nodules were identified, the studies included in these meta-analyses did not report the performance of PET-CT based on nodule size or on pre-test likelihood of malignancy, as categorised in clinical guidelines.

Evidence on CT surveillance was limited, with one study reporting diagnostic accuracy data. That found that volume doubling time and nodule volume had very high diagnostic accuracy to detect malignant nodules.

There was adequate evidence providing diagnostic accuracy estimates methods for CT-guided transthoracic needle biopsy. Better quality studies of radial probe endobronchial ultrasound (r-EBUS)-guided transbronchial lung biopsy may be needed, although they are probably less widely used than CT-guided biopsy.

Clinical impact of EarlyCDT Lung

No evidence was found on the clinical impact of using EarlyCDT Lung to diagnose pulmonary nodules. Instead, the EAG used simulation methods to investigate the possible impact of using EarlyCDT Lung. As the simulation was based on limited evidence, and required a number of strong assumptions to be made, its results should be treated as suggestive only.

Th simulation concluded that EarlyCDT Lung is unlikely to offer meaningful clinical improvement for low-risk nodules (0-10%), as adding EarlyCDT Lung to Brock risk appears to result in little change in diagnostic accuracy over using Brock risk alone. It appears to identify few additional genuinely malignant nodules and may lead to more false-positive results than true-positives.

At the 70% risk threshold, adding EarlyCDT Lung to Herder risk may improve sensitivity for only a small decline in specificity. Consequently, a large proportion of malignant nodules in the intermediate risk group (10%-70%) might be correctly identified by EarlyCDT Lung, and mostly reclassified to having a new risk of over 70%, with comparatively few false-positive reclassifications.

Cost-effectiveness reviews

The review of existing cost-effectiveness evidence identified two relevant studies. Neither of these was considered suitable to inform the current decision problem due to important differences, namely, in the patient population, the position and use of EarlyCDT Lung within the diagnostic pathway, and the diagnostic accuracy evidence used to inform it.

The additional reviews to support conceptualisation identified 8 diagnostic cost-effectiveness studies, and 34 screening studies; a sample of 9 screening models were reviewed. These reviews highlighted

that all evaluations relied on a common value mechanism of earlier diagnosis of lung cancer (at an earlier stage of disease). The reviews also identified structural assumptions and parameters estimates that could be used in alternative to those implemented in the EarlyCDT Lung cost-effectiveness studies.

Conceptualisation of cost-effectiveness model

The conceptualisation process identified three core components for a future cost-effectiveness assessment of EarlyCDT Lung: i) the characteristics of the subpopulations (reflecting the proposed positionings for EarlyCDT Lung in the current diagnostic pathway), ii) the way EarlyCDT Lung test results affect subsequent clinical management decisions, and iii) how changes in these decisions can affect outcomes.

There is limited evidence on the subpopulations of interest. Existing evidence, however, highlights that these are likely to differ in characteristics that drive value (such as prevalence of disease), and that there may be further heterogeneity (e.g., on outcomes).

The evidence on how EarlyCDT Lung test results are expected to affect subsequent management decisions indicates that this depends on the test's positioning, on nodule and patient characteristics (determining eligibility for subsequent management options), and the level of variation in clinical practice.

Changes in management decisions may affect clinical outcomes in two ways (two components of value). The first relates to short-term impacts (costs and adverse events) of escalating the current pathway to more interventional investigations/treatments (including the possibility of intervention on indolent malignant and benign nodules), and the potential for increased radiation exposure. The second relates to longer term health benefits and cost implications of earlier and/or increased detection (and treatment) of lung cancer. The evidence linkage mechanism for this component of value encompasses:

> i) the identification of differences in the time to diagnosis between current and proposed identification strategies, and mapping of these differences against likelihood or time to pre-clinical stage progression, to define the level of stage shift, and
>
> ii) the linking of the stage distributions, with and without stage shift, to expected long term outcomes conditional on disease stage.

There is little evidence on the time to diagnosis and the likelihood of stage progression under CT surveillance (and on heterogeneity on this), and on the potential for stage shift of EarlyCDT Lung. Linkage to health outcomes requires evidence on survival, HRQoL and costs conditional on disease stage at diagnosis. Our reviews identified UK-specific evidence on these components. Future cost-effectiveness models also need to consider: other determinants of outcomes (such as age or histology), primary tumour treatment, the need for adjustments for lead and length time bias (typically associated with stage shift mechanisms), and the adequacy of the data in reflecting contemporary treatments for lung cancer.

Conclusions

Implications for healthcare

The EAG concludes that the current evidence on EarlyCDT Lung is insufficient to determine its clinical value. This is due to the limited size of the relevant evidence base, and uncertainties as to whether current evidence generalises to the UK diagnostic pathway.

It appears that EarlyCDT Lung has poor diagnostic accuracy when used in isolation to diagnose pulmonary nodules, with low sensitivity to detect malignancy. It is therefore unclear what it can add to existing diagnostic methods, such as Brock and Herder risk assessment and the use of CT surveillance.

Based on results from the EAG's limited simulation study, EarlyCDT Lung may have little clinical benefit when diagnosing low-risk or smaller nodules, as it appears unlikely to appropriately change clinical management decisions. EarlyCDT Lung may possibly have clinical value when identifying malignancy in intermediate-risk nodules (10-70% risk), by correctly identifying high risk nodules that are malignant, and so might benefit from prompt excision.

There is no relevant evidence on the cost-effectiveness of EarlyCDT Lung and there is currently insufficient evidence to support explicit quantifications of the clinical and economic value of EarlyCDT Lung. We have identified key components and drivers of value that would need to be quantified in a future assessment of the clinical and economic value, and present considerations to support the conceptualisation of a future decision model.

Recommendations for research

Large, independent, prospective cohort studies, where EarlyCDT Lung is used in patients with identified pulmonary nodules are required. Patients should be diagnosed and manged in line with the

BTS diagnostic pathway. This will permit the estimation of the diagnostic accuracy of EarlyCDT Lung in isolation, and in combination with Brock and Herder risks. These studies should be used to validate, or update, the risk model proposed by Oncimmune.

These cohort studies should also assess the clinical impact of EarlyCDT Lung by reporting outcomes including:

- Impact on risk classification
- Change in clinical management
- Timing and tumour stage at detection and treatment of malignant nodules
- Avoidance of unnecessary CT or PET-CT scans
- Promotion of unnecessary PET-CT scans, biopsies or surgical excisions

Ideally, a randomised controlled trial should be performed, where patients with identified pulmonary nodules are randomised either to standard BTS management or to BTS management with EarlyCDT Lung included.

Currently, the broader evidence base on the whole BTS diagnostic pathway is limited. Large well-designed and UK-based prospective cohort studies are particularly needed to investigate the following:

- The diagnostic accuracy and clinical impact of using the Brock and Herder risk models
- The clinical consequences of CT surveillance, and
- How patient and nodule characteristics determine malignancy prevalence, eligibility for alternative clinical management options, likelihood and time to detection under CT surveillance, and patient outcomes.

A well-designed cost-effectiveness study is required, integrating emerging relevant evidence with the recommendations in this report to appropriately justify the value components considered and their translation into a relevant model structure.

PLAIN ENGLISH SUMMARY

People at risk of lung cancer sometimes undergo computed tomography (CT) scans of their lungs. These may identify lung nodules which could be cancerous. Currently CT scans of the lung nodules, or sometimes further Positron Emission Tomography (PET-CT) scans, are used to predict the risk that a nodule is cancerous.

EarlyCDT Lung is a blood test that detects substances, called autoantibodies, associated with having cancer. If the autoantibodies are detected the chance of a lung nodule being cancerous may be substantially increased. This test could help doctors make decisions about whether to treat immediately, carry out further tests, or monitor the nodule over time to see whether it grows or changes shape.

This project examined the evidence on the clinical value of the EarlyCDT Lung test. We reviewed all published studies of EarlyCDT and reanalysed the reported data. We found that there has been little research on EarlyCDT Lung in people with lung nodules (only five studies including 695 patients). This makes it difficult to draw any firm conclusions. The evidence suggests that EarlyCDT Lung may not be particularly effective at determining which lung nodules are cancerous, and may not improve diagnosis when compared to using CT and PET-CT scans. However, this is uncertain because the evidence is so limited.

This project also looked for evidence on the value for money of the EarlyCDT Lung test in detecting lung cancer, and found no relevant evidence. This means that the value for money of EarlyCDT Lung is largely unknown, and there is currently no good evidence to support further analyses on this. We therefore sought to summarise the information and analyses that would be needed to support a future assessment of the value for money of EarlyCDT Lung. We also made recommendations on what further studies may be important.

Table of Contents

Table of Tables

Table of Figures

List of abbreviations

NIHR	National Institute for Health Research
NHS	National Health Service
PET	Positron emission tomography
NICE	National Institute for Health and Clinical Excellence
CT	Computerised Tomography
MRI	Magnetic resonance imaging
EUS-FNA	Endoscopic ultrasound-guide fine-needle aspiration
EBUS-TBNA	Endobronchial ultrasound-guided transbronchial need aspiration
SCLC	Small cell lung cancer
NSCLC	Non-small cell lung cancer
SABR	Stereotactic ablative body radiotherapy
RFA	Radiofrequency ablation
BTS	British Thoracic Society
VATS	Video-assisted thorascopic surgery
MDT	Multi-disciplinary team
ELISA	Enzyme-linked immunosorbent assay
CE-IVD	CE-Invitro medical device
EAG	Evidence assessment group
ROC	Receiver operating characteristic
SF36	Short Form survery-36
EQ-5D	Euroquol-5 dimensions
WHO	World Health Organisation
AUC	Area under curve
QUADAS-2	Quality assessment tool of diagnostic accuracy studies
HSROC	Hierarchical summary receiver operating characteristic
PRISMA	Preferred reporting items for systematic review and meta-analyses
HIPAA	Health insurance portability and accountability authorisation
ECLS	Early diagnosis of lung cancer Scotland
PPV	Positive predictive value
CI	Confidence interval
PANAS	Positive and negative affect scale
LCWS	Lung cancer worry scale
IES	Impact of events scale
HAS	Health anxiety worry subscale

IPQR	Illness perception questionnaire-adapted for lung cancer
NR	Not reported
LR-	Negative Likelihood ratio
DOR	Diagnostic odds ratio
LR+	Positive Likelihood ratio
FDG	Fluorodeoxyglucose
SUV	Standardised uptake value
UKLS	UK Lung cancer screening
NPV	Negative predictive value
VDT	Volume doubling time
r-EBUS	Radial probe endobronchial ultrasound
PTN	Percutaneous transthoracic biopsy

ASSESSMENT GROUP REPORT

1 BACKGROUND AND DEFINITION OF THE DECISION PROBLEM

1.1 Lung cancer

Lung cancer is one of the most common types of cancer with around 47,000 diagnoses per year in the UK [1]. Lung cancer is commonly associated with smoking, being responsible for over 70% of cases. Other causes of cancer include passive smoking and exposure to asbestos or other carcinogenic chemicals.

1.1.1 Diagnosis of lung cancer

Lung cancer is often diagnosed later and at a more advanced stage than for other cancers. Early detection is critical for improving outcomes. Diagnosis of lung cancer requires more than one investigation. Initial investigations may involve assessment of clinical symptoms and signs to exclude other illnesses, such as chest infections.

NICE guidance on diagnosis and management lung cancer 2019 makes several recommendations that optimise the diagnostic pathway and allow flexibility for managing symptoms of lung cancer in a range of people [2]. The guideline recommends that patients with suspected lung cancer should be urgently referred for a chest X-ray. If the results suggest lung cancer, a contrast-enhanced CT scan of the chest, upper abdomen and lower neck is performed.

Further investigations to confirm a diagnosis and to provide information on the stage of the disease are then carried out. These investigations generally include a biopsy for histological confirmation and subtyping but may also include positron emission tomography-computed tomography (PET-CT). Other methods that can diagnose and stage the disease are MRI, endobronchial ultrasound-guided transbronchial needle aspiration (EBUS-TBNA) and endoscopic ultrasound-guide fine-needle aspiration (EUS-FNA) [2]. This helps with diagnosis and choosing the best treatment.

1.1.2 Diagnostic pathway for pulmonary nodules

Pulmonary nodules are small growths in the lung, often found when having a chest X-ray or CT scan; for example, when performing a CT scan for conditions unrelated to cancer (incidental findings), when patients are referred to the diagnostic pathway from symptoms, or as part of lung cancer screening. They may be malignant or benign and in the UK are generally managed in accordance with the British Thoracic Society Guidelines for the investigation and management of pulmonary nodules (2015) [3]. In America, the Fleischner Society Guidelines for management of solid nodules (2005) [4] are

widely used, but these are not often followed in the UK. Other guidelines, such as those of the American College of Chest Physicians [5], are also available. Figure 1 provides a recommended pathway for the initial approach to solid pulmonary nodules.

For nodules smaller than 5 mm in diameter (or 80 mm^3 in volume), the British Thoracic Society recommend that people should be discharged with no follow up. People with nodules of 5 to 8mm diameter, or under 300 mm^3 in volume, which are expected to have lower than 10% risk of malignancy, are offered CT surveillance. This involves repeat scanning at 3 months, 1 year, and sometimes 2 years to assess nodule volume doubling time. The frequency and duration of CT scans is determined by nodule size and characteristics.

For larger nodules (over 8mm in diameter) the Brock model is used to assess risk of malignancy. If risk is low (<10%) people will be offered CT surveillance. For pulmonary nodules at above 10% risk after Brock model assessment, PET-CT is recommended, and the nodule risk is then recalculated based on the Herder model. The Herder model predicts the risk of malignancy in solid pulmonary nodules using patient characteristics, nodules characteristics, and the degree of F-fluorodeoxyglucose uptake on PET-CT. [6]

For people with 10-70% risk of malignancy using the Herder model, image guided biopsy, excision biopsy or CT surveillance guided by individual risk and patient preference is used. Image-guided percutaneous lung biopsy is recommended in patients with peripheral pulmonary lesions. Non-imaging tests such as bronchoscopy (augmented using either radial endobronchial ultrasound, fluoroscopy or electromagnetic navigation) can also be performed for pulmonary nodules with bronchus sign present on CT.

People with risk over 70% are considered for excision or non-surgical treatment (see Figure 1 for more information).

1.1.2.1 Excision and surgery

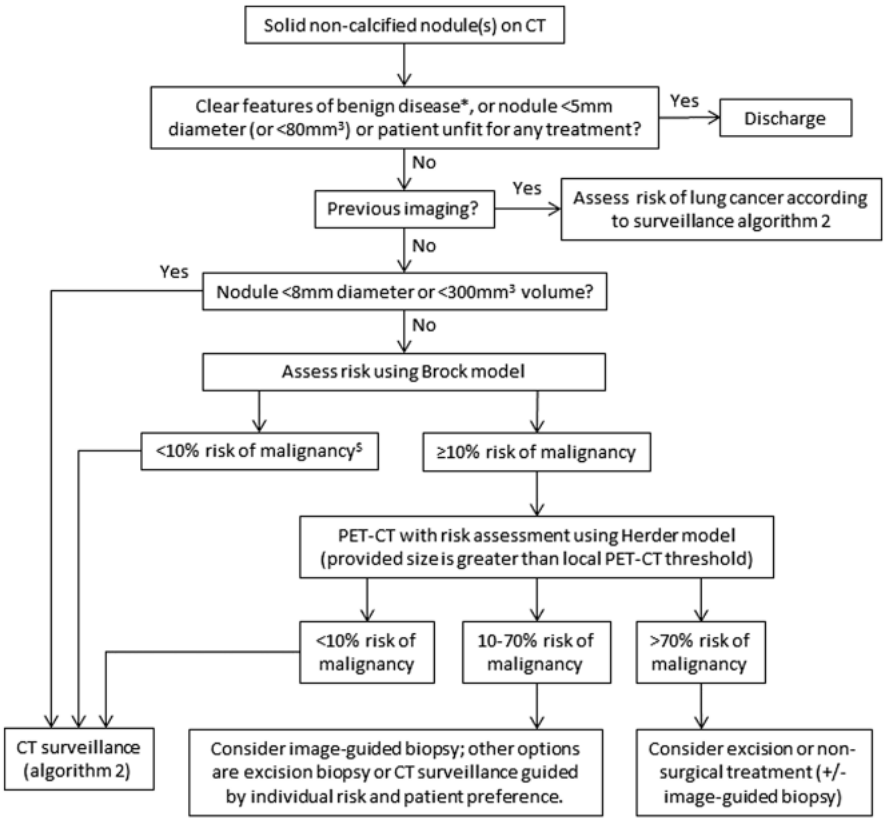
The treatment of choice in early lung cancer is excision, with non-surgical treatment only considered in people who are not fit for surgery.

Excision of pulmonary nodules is performed in three situations: either where there is confirmed malignancy from preoperative biopsy; or where a nodule is considered of sufficiently high risk to merit excision with no preoperative biopsy or after a negative biopsy; or where biopsy had an indeterminate result.

If malignancy is not confirmed by image-guided biopsy, nodules in the lung periphery are suitable for wedge resection with intraoperative frozen section pathological analysis. This approach has been shown to present high sensitivity and specificity with a definitive diagnosis achieved in all cases, and low rates of morbidity and mortality in relation to lobectomy from limiting the extent of lung resection for benign disease.

For surgical excision of a pulmonary nodule, the BTS guidelines prefer video-assisted thoracoscopic surgery (VATS) to an open approach (thoracotomy). In what concerns the extent of the lung resection, lobectomy should be offered to patients fit enough to undergo the procedure, as definitive management of a confirmed lung cancer pulmonary nodule (during the same anaesthetic procedure if confirmed during wedge resection). Anatomical segmentectomy (sublobar resection) should be considered if the patient is unfit for lobectomy, and as diagnostic for nodules <2 cm in diameter without nodal disease when there has been no pathological confirmation.

Figure 1 Initial approach to solid pulmonary nodules (British Thoracic Society guidelines 2015)



1.2 Population and relevant subgroups

The population of interest is all persons with solid non-calcified pulmonary nodules identified by CT scanning, whether received for conditions unrelated to lung cancer, as part of a cancer diagnosis procedure for people with possible lung cancer symptoms, or as part of a lung cancer screening programme. Specifically, the assessment will examine:

1. People with a nodule of 5-8mm in diameter or 80-300mm^3 in volume
2. People with <10% risk of malignancy using the Brock model after initial CT scan or using the Herder model after PET-CT scan
3. People with 10-70% risk of malignancy using the Brock model, or the Herder model (after PET-CT scan)

People with other cancers, or who have had a cancer diagnosis in the past five years, are excluded from consideration: EarlyCDT Lung is not recommended for such persons.

Table 1 shows the possible distribution of patients at different parts of the BTS pathway, taken from a study by Al Ameri et al of 186 individuals which match the population of interest for this assessment This study suggested that the majority of patients presenting with incidentally detected nodules are classed as small (nodules between 5-8mm diameter) or low risk. These are assigned to CT surveillance, which suggests a large burden on the health system from the multiple follow-up CT scans. This evidence also suggests there is a meaningful proportion of cancers detected at metastatic disease across all risk groups.

Table 1 Distribution of patients across the BTS pathway (after Al-Ameri 2015)

Risk groups (as defined by the BTS pathway)	% (n)*	Prevalence, %		
		Overall	Primary cancer	Metastatic cancer
Low risk, people with nodules 5-8mm in diameter, or with <10% risk of malignancy using Brock and Herder (referred to CT surveillance)	57.0% (106)	6%	2%	4%
Intermediate risk, people with 10-70% risk of malignancy using the Herder model	31.2% (58)	64%	55%	9%
High risk, people with >70% risk of malignancy using the Herder model	11.8% (22)	91%	81%	10%
Total	100% (186)	34.1%	27.8%	6.3%

* % of those with nodules > 5mm diameter

In all populations, patients would receive an EarlyCDT Lung test and proceed to excision or surgery if deemed to be at high risk of malignancy (>70%). At lower risk of malignancy (<70%) patients would go on to CT surveillance, or possibly biopsy or excision for patients at intermediate risk (10-70%).

The protocol specified that the key subgroups of interest were the different reasons for receiving an initial CT scan (patients with symptoms, incidental finding when scanning for other conditions, or as part of a cancer screening programme). However, no data on these subgroups was identified, so they could not be investigated.

1.3 Diagnostic technologies under assessment

1.3.1 Early CDT Lung

EarlyCDT Lung is a blood test that can be used to assess the malignancy risk of people at risk of lung cancer. The test can, in principle, be used on any at-risk person; this assessment will consider its use in persons with solid pulmonary nodules found by chest CT scan or X-ray.[7-9] Incidental finding of pulmonary nodules in asymptomatic individuals, when performing CT scans for other medical purposes, or during lung cancer screening, is an increasingly common clinical dilemma encountered by lung cancer clinicians. EarlyCDT Lung could be used as part of the standard diagnostic pathway for early detection of lung cancer, where it might result in treatment being offered earlier, giving improved patient outcomes.

EarlyCDT Lung uses a standard enzyme-linked immunosorbent assay (ELISA) method. It is manufactured by Oncimmune and is available as a CE-IVD marked kit. It was launched commercially in November 2010 with physicians in routine practice across the USA ordering the test on behalf of their patients.[10] The test measures the presence of autoantibodies to a panel of 7 lung cancer associated antigens (p53, NY-ESO-1, CAGE, GBU4-5, HuD, MAGE A4 and SOX2). [7] A blood sample is considered positive when at least one of the 7 autoantibodies is elevated above a pre-determined cut-off (Table 2). Elevated levels of these autoantibodies may indicate current (or past) malignant disease. The thresholds were set to give a high test specificity with the aim of reducing false-positive results that would lead to unnecessary and potentially invasive diagnostic procedures. The EarlyCDT Lung test results are interpreted by skilled medical professionals in combination with other clinical information. In particular, it is suggested that its results be used to modify the risk of malignancy estimated by existing nodule risk calculators, including the Brock model and the Herder model. [11, 12]

Table 2 Recommended cut-offs for autoantibodies measured using EarlyCDT Lung

Autoantibody		Low cut-off value		High cut-off value	
CAGE		4.25		5.27	
GBU4 5	**No significant level of autoantibodies detected**	4.36	**Moderate level result**	5.92	**High level result**
NYESO 1		3.02		4.27	
p53		5.79		6.47	
SOX2		5.48		5.58	
MAGE A4		6.19		7.94	
HuD		7.31		8.15	

Oncimmune have described EarlyCDT Lung as a "rule-in" test to help identify pulmonary nodules that may benefit from earlier diagnosis and treatment. Results of EarlyCDT Lung tests are reported as one of three options:

- No significant levels of autoantibodies detected
    - (if no autoantibody is above the low cut-off level)
- Positive-moderate
    - (if at least one autoantibody is above the low cut-off level, but below the high-cut-off level)
- Positive-high
    - (if at least one autoantibody is above the high cut-off level)

A patient will have a pre-test risk of lung cancer predicted by their sex, age, smoking history, and other risk factors alone, calculated by the Brock (or Swensen/Mayo) nodule malignancy risk calculator. If a person is being assessed after PET-CT scan their risk may be assessed using the Herder malignancy risk tool.

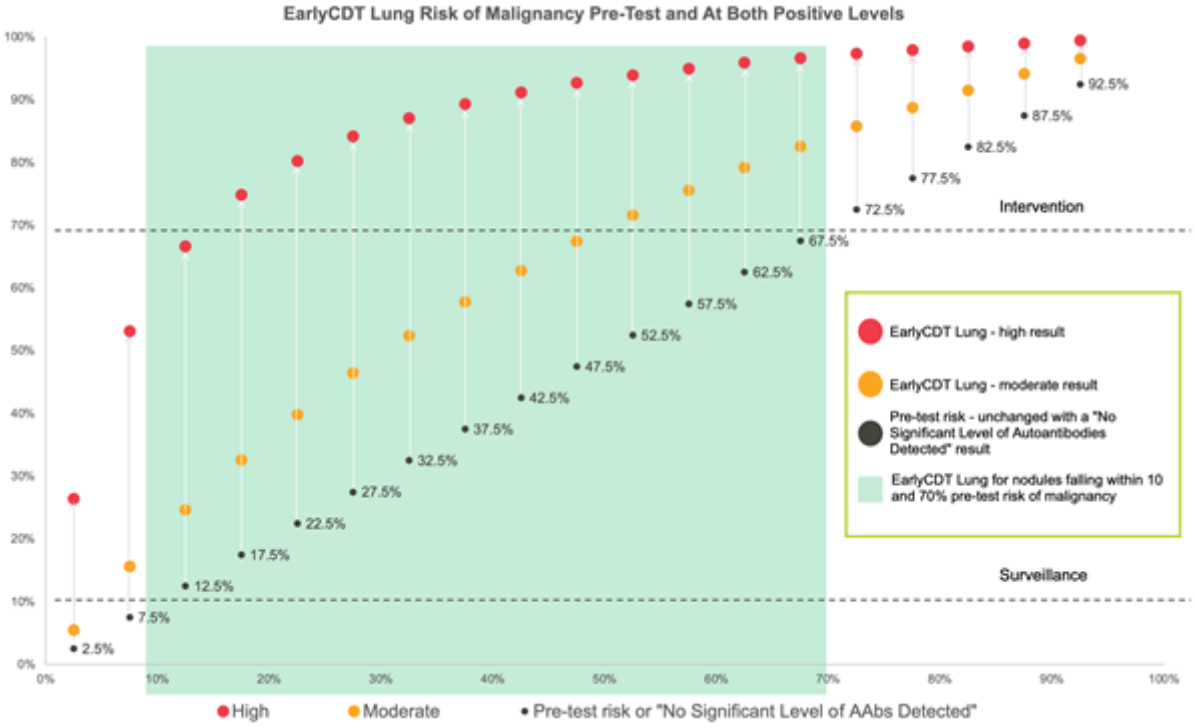
Oncimmune proposes that the EarlyCDT Lung test result is used to update these estimated risks of malignancy. For people who test negative with EarlyCDT Lung, Oncimmune recommends that the estimated risk is left unchanged from the pre-test risk – in this way defining this test as a 'rule-in' test. Statistically, a patient with a negative test result should see their risk scores downgraded, but this is not proposed for this assessment. Clinical management in these individuals would then proceed in line with the pre-test risk.

A positive-moderate result would lead to a moderate increase in the chance of malignancy from the pre-test risk. If the increase in risk is large enough it might suggest that further diagnostic testing is needed, such as image-guided biopsy. A positive-high result would lead to a larger increase in the

chance of malignancy from the pre-test risk. This might suggest that further diagnostic testing is needed, or if the new risk estimate is sufficiently high, that the person should proceed directly to surgical resection of the nodules.

Oncimmune have produced a graph detailing how the pre-test risk could be modified given a positive-moderate or positive-high EarlyCDT Lung test result (Figure 2). The calculation of post-test mortality risk from the baseline risk obtained from the Swensen/Mayo calculator and the EarlyCDT Lung test result is described in Healey et al. (2017). [10] Oncimmune proposes applying this calculation to pre-test risks derived with both Brock and Herder models.

Figure 2 Impact of EarlyCDT Lung on lung cancer risk assessment



The EarlyCDT Lung test should not be used in people with a previous history of cancer of any type, except for basal cell carcinoma, as other cancers may lead to elevated levels of autoantibodies, and hence to false-positive results. It should also not be used in people known to have diseases that result in an elevated level of serum total protein, for example, myeloma, amyloidosis, and monoclonal gammopathy of undetermined significance.

As far as the EAG can determine, the EarlyCDT Lung test is not currently in regular use in the UK for the assessment of pulmonary nodules. It has been used in a large-scale trial in Scotland to investigate whether it can be used as part of a lung cancer screening programme [13].

1.3.1.1 Cost of EarlyCDT Lung testing

The cost of EarlyCDT Lung testing to the NHS includes the costs of i) the test, ii) consumables required to process the test, iii) test administration iv) training needed to process/administer the test, and v) costs of delivering test results to individuals.

In response to a request for information by NICE, Oncimmune, provided an estimate of the cost per EarlyCDT kit of £600 (excluding VAT). According to the company each kit can run up to 10 patient samples. Therefore, assuming that patient throughput is sufficient to always ensure simultaneous processing of 10 patient samples, and that there are no test failures, the cost of test per patient will be £60. When determining this element of cost in future assessments, patient throughput will have to be considered, as throughput may vary according across NHS trusts and will be a determinant of cost. Furthermore, the need to "batch" tests for processing may lead to delays to diagnosis in trusts with low throughput who may have to wait for a sufficient number of samples before processing a full batch. The extent of this delay is unknown, but it may reduce some of the benefit from early diagnosis with EarlyCDT Lung if too prolonged.The company also provided an estimate of the costs of consumables required to process each EarlyCDT Lung kit in their laboratories, but notes that costs may differ in a NHS setting (the company suggests that these costs may be lower to the NHS without supporting evidence). The costs of consumables to process one EarlyCDT Lung as provided by the company are shown in Table 3; it suggests a cost per kit of £7.13 and a cost per test of £0.71.

Table 3 Cost of consumables for processing one EarlyCDT Lung (up to 10 patient samples)

2-plate format for running 10 tests						
Item	**quantity used**	**Cost per case £**	**pack size**	**unit cost**	**cost per kit (£)**	**cost per test (£)**
5ml pipette tips	2	63.21	1000	0.06	0.13	0.01
serum dilution tubes	10	99.76	700	0.14	1.43	0.14
20ul pipette tips	10	77	960	0.08	0.80	0.08
1000ul pipette tips	13	37.78	768	0.05	0.64	0.06
1200ul pipette tips	24	84	960	0.09	2.10	0.21
dH20 (L)	1	7.32	5	1.46	1.46	0.15
Falcon tube (50mL)	1	56.04	500	0.11	0.11	0.01
reagent troughs	3	15.26	100	0.15	0.46	0.05
Total cost:					**7.13**	**0.71**

Table adapted from company's response to a request for information

The EAG notes that these costs are likely to vary across NHS trusts, as local procurement arrangements may result on variation of the costs per item for each consumable. The patient throughput is less likely to affect the calculations of cost per patient, assuming that most of the listed consumables are routinely used materials in NHS laboratories.

The company did not report any estimates for the remaining elements of costs (i.e., the costs of test administration, training needed to process/administer the test, and costs of delivering test results to individuals).

According to EarlyCDT Lung's instructions for use (as submitted by Oncimmune), the test requires a blood sample (serum or plasma) to be collected. Therefore, the test administration cost should reflect the NHS staff time required to collect the blood sample. This cost may vary depending on local protocols, and whether these impose additional contacts with health care professionals. For example, if the test is to be administered only to patients at risk level below 70% on the Herder score, the patient may require one additional contact with the health service after the PET-CT scan (see Section 1.5) to have their blood sample collected, and this cost will only be incurred by patients with a risk score below 70%. Alternatively, local protocols might require a sample to be collected for all patients when they receive the PET-CT; this may allow for some efficiency gains if blood collection can be fitted within the work-up. However, it means that the cost of collecting the blood is incurred for all patients regardless of their Herder score, and that the blood sample needs to be stored until the Herder score is available for all patients. Another option to avoid storing unnecessary blood samples, is to test all samples EarlyCDT Lung regadless of the patients risk score; this would imply all patients incur the cost of the test, as well as the cost of collecting a blood sample.

The company did not provide information on the training requirements needed to process and interpret the test, but these should also need to be considered. These costs would need to reflect the cost of laboratory staff time to learn how to run the test and use the associated software to obtain a result. Similarly, the cost of clinicians time to learn how to interpret the results should also be included.

Finally, a cost may have to be included to reflect additional time for the clinician to interpret the test results (which may be negligible in the context of the diagnostic workup) and, more importantly, any additional contacts between the patient and the health care system to deliver the results of the test. The BTS guidelines recommend offering "*patients the choice of seeing a lung cancer nurse specialist where the probability of malignancy is high or when patients are anxious about the possibility of having lung cancer"*.[3] So it may be appropriate to include the cost of an appointment with a specialist nurse for some of the patients. While this cost may also be incurred in strategies without EarlyCDT

Lung, the number of patients incurring the cost will vary across strategies and it should, therefore, be considered.

1.3.2 Other technologies

This report does not consider other novel technologies for the diagnosis of lung cancer, including other autoantibody tests or lung cancer risk assessment tools. At present, no suitable alternative technologies have the relevant approval for use in the UK.

1.4 Comparators

The overall comparator was the current BTS recommended diagnostic pathway for pulmonary nodules without EarlyCDT Lung (Figure 1). Specifically, this included diagnosis and management of nodules using:

1. The Brock model
2. The Herder model (after PET-CT)
3. No risk assessment (for nodules between 5-8mm in diameter or 80-300mm^3 in volume)

In order to fully interpret the clinical and economic impact of using EarlyCDT Lung the diagnostic accuracy and clinical effectiveness of the following specific parts of the diagnostic pathway were also investigated:

4. CT surveillance (for small or low-risk nodules)
5. PET-CT scans (for intermediate risk nodules)
6. Biopsy of suspicious nodules (for high-risk nodules)

1.5 Place of the intervention in the care pathway

Lung cancer is often diagnosed at a more advanced stage than other common cancers. National Cancer Registration and Analysis Service data show that almost half of all lung cancers are diagnosed at stage 4. Late diagnosis, where curative treatment is not possible, is a contributing factor to poor survival rates for people with lung cancer. Early detection is key to improving outcomes.

The proposed position of EarlyCDT Lung test within the current British Thoracic Society pathway for solid pulmonary nodules (British Thoracic Society guidelines 2015) is shown in Figure 3. This pathway includes an option where PET-CT scans are not available. Clinical opinions received at scoping, and during the project, suggested that lack of access to PET-CT is not of concern for the NHS. This assessment will therefore only consider the part of the pathway where PET-CT is available.

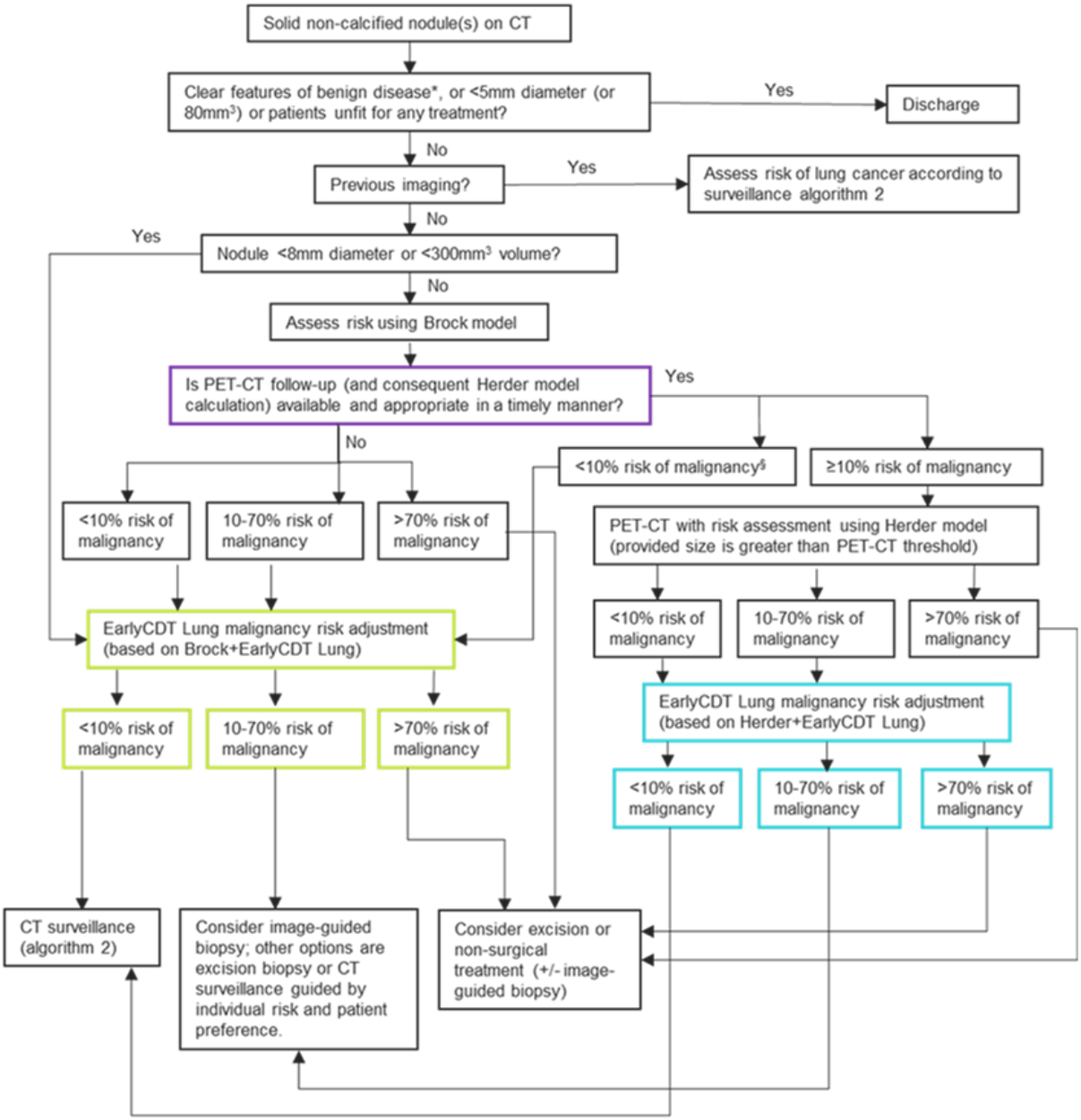
The position of EarlyCDT Lung has been stated to be after the first CT scan, or post PET-CT when the result suggests intermediate risk. EarlyCDT Lung could be used to assess people with nodules

<8mm diameter or 300mm^3 volume and those with <10% risk of malignancy after using the Brock model. The test could also be used for people with 10-70% risk of malignancy after using either the Brock or the Herder models. If the EarlyCDT Lung test is positive, the malignancy risk is increased and people with a post-test risk of greater than 70% could then be moved into the intervention pathway immediately, without the delay caused by CT surveillance, or further diagnostic testing.

This assessment will consider the following specific locations in the diagnostic pathway where EarlyCDT Lung could be used, the feasibility and relevance of the proposed placements will be established based on clinical advice:

1. For people with nodules 5-8mm in diameter or 80-300mm^3 in volume
2. In combination with CT scan and Brock model, in people with nodules >8 mm in diameter that have <10% risk of malignancy using the Brock model after initial CT scan
3. In combination with PET-CT scan and Herder model, in people with nodules >8 mm in diameter that have <10% risk of malignancy using the Herder model after PET-CT scan
4. In combination with CT scan and Brock model, in people with nodules >8 mm in diameter that have 10-70% risk of malignancy using the Brock model (with EarlyCDT Lung preceding PET-CT)
5. In combination with PET-CT scan and Herder model, in people with nodules >8 mm in diameter that have 10-70% risk of malignancy using the Herder model, after PET-CT scan

Figure 3 Proposed position of EarlyCDT Lung within the current BTS pathway for lung cancer



Key

* e.g. hamartoma, typical peri-fissural nodule

§ Consider PET-CT for larger nodules in young patients with low risk by Brock score as this score was developed in screening cohort (50-75 years) so performance in younger patients unproven.

1.5.1 Action after risk assessment

Under the current diagnostic pathway (Figure 1) persons with small nodules or a low malignancy risk (<10%) are offered CT surveillance, with regular CT scans to check for growth of the nodules. Persons with high-risk nodules (>70%) proceed directly to excision or treatment, if suitable, with a biopsy for confirmation, where required. For persons with intermediate risk (10-70%) there are a

wider range of options. These include: image guided biopsy or excision biopsy, or CT surveillance. The exact choice of approach will depend on the estimated risk, clinical opinion, and patient preference.

EarlyCDT Lung is proposed to update the individual's risk, but it is currently unclear if or how clinical decision making, conditional on the updated risk score, would be altered. Clinical advice received suggests there may be some uncertainty or difference of opinion. For example, whether patients with small nodules but a positive EarlyCDT Lung test could undergo biopsy, as the nodule may be too small to biopsy effectively; or what level of risk to change from CT surveillance to image-guided biopsy. This assessment investigated the following general pathway after EarlyCDT Lung assessment.

- For small or low risk nodules where risk is below 10% risk after EarlyCDT Lung
    - Offer CT surveillance in accordance with standard pathway
- For small or low risk nodules where risk increases to 10%-70% risk after EarlyCDT Lung
    - Consider PET-CT scan provided the nodule is above the size threshold for PET-CT
    - Consider image-guided biopsy (may not be feasible for smaller nodules)
    - Offer CT surveillance (possibly at higher frequency for small nodules where PET-CT or image-guided biopsy not helpful/possible respectively)
    - Consider excision biopsy
- For small or low risk nodules where risk increases to over 70% risk after EarlyCDT Lung
    - This may not be possible given working of risk algorithm
    - All patients being considered for surgery or treatment would need PET-CT for staging
    - Image-guided biopsy prior to surgery/treatment may be considered
- For intermediate risk nodules still at 10%-70% risk after EarlyCDT Lung
    - Proceed as for standard pathway, although choice of action may be influenced by any change in estimated risk within the 10%-70% spectrum (e.g. more likely to proceed to biopsy at higher risk)
- For intermediate risk nodules where risk increases to over 70% risk after EarlyCDT Lung
    - Proceed directly to excision or treatment
    - All patients being considered for surgery or treatment would need PET-CT for staging
    - Image-guided biopsy prior to surgery/treatment may be considered

1.6 Outcomes

Below is a list of all key outcomes judged to be relevant to the assessment of the clinical and cost-effectiveness of EarlyCDT Lung, and the general diagnostic pathway for pulmonary nodules. This represents a comprehensive list of outcomes listed in the protocol. Owing to the limited nature of the published literature, particularly for EarlyCDT Lung, many of these outcomes could only be evaluated

using indirect evidence (such as data from lung cancer screening studies), or could not be formally assessed.

- Diagnostic accuracy
    - Sensitivity, specificity, positive and negative predictive values, diagnostic likelihood ratios, areas under ROC curves
    - For EarlyCDT Lung in isolation and in combination with Brock and Herder models
- Short-term clinical outcomes
    - Impact of test on risk classification
    - Impact on clinical decisions relating to diagnostic or treatment pathway
    - Further tests used
        - Including PET-CT and image-guided or excision biopsy
    - Adverse events during or after testing
- Longer-term clinical outcomes
    - Lung cancer mortality
    - Lung cancer related morbidity
    - Morbidity associated with other diagnostic tests or procedures
    - Overall and disease-free survival
- Patient-focussed outcomes
    - Health-related quality of life
        - SF36, EQ-5D
    - Impact on anxiety and cancer concern
        - False-positive tests
        - Unnecessary biopsies or other procedures
        - Overdiagnosis of tumours not requiring immediate treatment
        - Delay in diagnosing treatable cancers
        - Understanding and communication of test results
- Implementation of test
    - Time to obtain results
    - Laboratory capacity
    - Training requirements
    - Clinical variation in interpreting and using results

2 ASSESSMENT DESIGN

2.1 Objectives

The aim of the project was to appraise existing evidence on the potential clinical and cost-effectiveness of the EarlyCDT Lung test for lung cancer risk classification of solid pulmonary nodules, and to develop a conceptual economic model to provide a common understanding of the evidence requirements and evidence linkages required to undertake a robust cost-effectiveness analysis. To achieve this, the following objectives were set:

Clinical effectiveness

- To perform a systematic review and, if feasible, a meta-analysis of the diagnostic accuracy of EarlyCDT Lung for lung cancer risk classification of solid pulmonary nodules.
- To perform a narrative systematic review of the clinical impact and practical implementation of using the EarlyCDT Lung test.
- To perform a scoping review of the evidence on EarlyCDT Lung for uses outside the specified diagnostic pathway (e.g. as a lung cancer screening tool), where this will inform the overall review.

Cost-effectiveness

- To perform a systematic review of published cost-effectiveness studies of EarlyCDT Lung for lung cancer risk classification of solid pulmonary nodules.
- To review cost-effectiveness models for other surveillance and diagnostic strategies for the identification of malignancy in solid pulmonary nodules, and cost-effectiveness models of screening strategies for lung cancer.
- To conceptualise a decision model structure to provide a common understanding of how the cost-effectiveness of EarlyCDT Lung for lung cancer risk classification of solid pulmonary nodules in the different positions of the diagnostic pathway proposed for the technology can be quantified.
- To scope existing evidence that could support the implementation of the conceptualised decision model, highlighting key evidential and structural uncertainties.

The objectives for this assessment were set out in the development of the protocol, which acknowledged that the existing published evidence base on EarlyCDT Lung was too small to allow a full assessment of the clinical and economic value of the test. This assessment was therefore restricted

to review the extent of the existing evidence and provide a common understanding of the evidence requirements and evidence linkages required for a full assessment of the value of EarlyCDT Lung to the NHS. The EAG was therefore not requested to develop and implement a *de novo* decision analytic model.

2.2 Systematic review of diagnostic accuracy and clinical effectiveness

The systematic review was conducted following the general principles recommended in CRD's guidance and is reported in accordance with the PRISMA statement [14, 15].

2.2.1 Literature searching

The aim of the literature search was to systematically identify all published and unpublished studies of the EarlyCDT Lung test.

An Information Specialist (MH) designed the search strategy in Ovid MEDLINE in consultation with the research team. The strategy consisted of a set of terms for the named technology EarlyCDT, with a further section focusing on terms for autoantibodies for detecting lung cancer or pulmonary nodules. Text word searches for terms appearing in the title and abstracts of database records were included in the strategy alongside searches of relevant subject headings. Date, language and study design limits were not applied. The final MEDLINE strategy was adapted for use in all resources searched.

The searches were carried out on 8th March 2021. The following databases were searched: MEDLINE (including: Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE Daily and Ovid MEDLINE), Embase, Cochrane Central Register of Controlled Trials (CENTRAL), Science Citation Index, EconLit, Cochrane Database of Systematic Reviews (CDSR), Database of Abstracts of Reviews of Effects (DARE), Health Technology Assessment (HTA) Database, NHS Economic Evaluation Database (NHS EED), and the International Health Technology Assessment Database.

In addition, the following resources were searched for on-going, unpublished or grey literature: ClinicalTrials.gov, EU Clinical Trials Register, Conference Proceedings Citation Index: Science, Proquest Dissertations and Theses A&I, Open Access Theses and Dissertations, and the International Prospective Register of Systematic Reviews (PROSPERO). The WHO International Clinical Trials Registry Platform portal was due to be searched however was not available during March 2021.

Search results were imported into EndNote 20 (Clarivate Analytics, US) and deduplicated. All search strategies are presented in full in Appendix 1. Reference lists of relevant reviews and studies were

scanned in order to identify additional potentially relevant reports. Forward citation searching of Science Citation Index was also used to identify relevant papers that cited key included papers.

2.2.2 Additional literature searching

In order to identify and appraise existing evidence on the clinical and cost-effectiveness of Early CDT Lung, and inform the conceptualisation of a decision model, it is anticipated that sources of evidence on the diagnosis, management and treatment of pulmonary nodules will be required, beyond that reported in the literature on EarlyCDT Lung.

Focussed and pragmatic searches of the databases were performed to identify literature on the diagnostic accuracy, clinical impact and cost-effectiveness of all the identified comparator technologies, specifically:

- Brock and Herder models
- CT screening and surveillance
- PET-CT scans
- Biopsy

Within the library of potentially relevant papers, key word searching was used to identify papers on the comparators listed above. Screening focussed on identifying systematic reviews in these areas. If systematic reviews were not available, trials and cohort studies of relevance to UK practice were identified.

2.2.3 Study selection

Two reviewers independently screened all titles and abstracts. Full papers of any titles and abstracts of potentially relevant papers and conference abstracts were obtained where possible, and the relevance of each study assessed independently by two reviewers according to the criteria below. Any disagreements were resolved by consensus or, where necessary, by consulting a third reviewer.

The following eligibility criteria were used to identify relevant studies:

2.2.3.1 Participants

Persons with solid non-calcified pulmonary nodules identified by CT scanning, who may be eligible for further screening or diagnostic testing, including using the EarlyCDT Lung test.

Subpopulations were people with:

1. nodules between 5-8mm in diameter or 80-300mm^3 in volume

2. nodules over 8mm in diameter and over 300mm^3 in volume with a risk of malignancy estimated to be under 10% (using either Brock or Herder model)
3. nodules over 8mm in diameter and over 300mm^3 in volume with a risk of malignancy estimated to be between 10% and 70% (using either Brock or Herder model)

Persons who have had a previous cancer diagnosis were excluded. Persons with a malignancy risk above 70% (before EarlyCDT Lung test) were also excluded, as they are recommended to proceed directly to surgical excision, and would not benefit from further testing.

2.2.3.2 Interventions

The EarlyCDT Lung test. The test was considered in three possible locations in the diagnostic pathway:

1. In isolation, for nodules between 5-8mm in diameter or under 300mm^3 in volume
2. In combination with the Brock test, where the Brock test suggests a malignancy risk of <10%
3. In combination with the Brock test, and/or Herder test after PET-CT scan, where an intermediate malignancy risk (10 – 70%) is estimated.

No other interventions were considered

2.2.3.3 Comparators

As stated in Section 1.4, the broad comparator was diagnosis and management of pulmonary nodules using current BTS guidelines (as in Figure 1). Specifically, this included diagnosis and management of nodules using:

7. The Brock model
8. The Herder model (after PET-CT)
9. No risk assessment (for nodules between 5-8mm in diameter or 80-300mm^3 in volume)

2.2.3.4 Reference standard

Two types of reference standard were eligible. Firstly, a confirmed diagnosis of a malignant or benign tumour by image-guided biopsy, excision biopsy or surgical resection. Secondly, the results of follow up visits for confirming the absence of malignancy; confirmed stable nodule volume after one year, or stable diameter after two years, were deemed to be the most appropriate durations.

2.2.3.5 Outcomes

Due to data limitations, outcome analysed were largely limited to diagnostic accuracy measures (sensitivity, specificity, area under ROC curve (AUC)), with some limited investigation of changes to risk classification. A full list of outcomes of interest is given in Section 1.6.

2.2.3.6 Study designs

Due to the anticipated small number of studies and publications likely to be eligible, all study designs were included, provided they reported evidence on the outcomes listed in Section 1.6.

All forms of evidence were considered, including both quantitative data and qualitative evidence.

2.2.4 Data extraction

Data on study and patient characteristics and results were extracted by one reviewer using a standardised data extraction form and independently checked by a second reviewer. Discrepancies were resolved by discussion, with involvement of a third reviewer where necessary. Where feasible, data were also electronically extracted from figures presented in publications.

Data from relevant studies with multiple publications were extracted and reported as a single study. The most recent or most complete publication was used in situations where we could not exclude the possibility of overlapping populations.

2.2.5 Quality assessment strategy

The quality of the diagnostic accuracy studies was assessed using the QUADAS-2 tool (Quality Assessment tool of Diagnostic Accuracy Studies), modified as necessary to incorporate review-specific issues. QUADAS-2 evaluates both risk of bias and study applicability to the review question. The Cochrane risk of bias tool was used to assess clinical trials.

The quality assessments were performed by one reviewer and independently checked by a second reviewer. Disagreements were resolved through consensus, and where necessary, by consulting a third reviewer.

2.2.6 Synthesis

The literature on the EarlyCDT Lung test, and on comparator technologies was small and largely insufficient to perform meta-analyses. Where sufficient clinically and statistically homogenous data were available, data were pooled using appropriate meta-analytic techniques (see Section 2.2.6.1). However, a narrative approach to synthesis was required for most of the comparators, with the results of data extraction being presented in structured tables, and plotted in figures where feasible.

2.2.6.1 Meta-analysis and narrative synthesis of diagnostic accuracy

Using extracted diagnostic accuracy data from 2 x 2 tables, or reported diagnostic accuracy results, estimates of sensitivity and specificity were calculated and presented on forest plots and in the

receiver operating characteristic (ROC) space to examine the variability in diagnostic test accuracy within and between studies. Positive and negative predictive values were calculated.

Where three or more studies were available the hierarchical summary ROC (HSROC) model was fitted to produce summary meta-analysis estimates of diagnostic accuracy and summary ROC curves. Univariate random-effects meta-analyses of diagnostic outcomes (sensitivity, specificity, diagnostic odds ratios and area under ROC curves (AUC)) were also performed where the HSROC model could not be fitted.

2.2.6.2 Synthesis of clinical outcomes, patient-focussed outcomes and implementation evidence

Data on outcomes other than diagnostic accuracy were rarely reported. Narrative synthesis was used where feasible, by comparing the tabulated results across studies to identify broad evidence of effectiveness.

Data on diagnostic accuracy for EarlyCDT Lung were combined with data on lung cancer prevalence and nodule risk based on the Brock and Herder models in order to simulate the potential clinical impact of using EarlyCDT Lung, in terms of changes in diagnostic accuracy and diagnostic pathway. For full details see Section 3.3.1.

2.2.6.3 Investigation of heterogeneity and subgroup analyses

For diagnostic accuracy data, we visually inspected the forest plots and ROC space to check for heterogeneity between study results. Where data permitted, subgroup analyses were used, by performing meta-analyses in defined subgroups of studies.

2.2.6.4 Sensitivity analyses

It was our intention to carry out sensitivity analyses to explore the robustness of the results according to study quality based on QUADAS-2 domain results (for example, by excluding studies with high risk of incorporation bias) and study design (for example, in-procedure versus retrospective evaluation of index test results). However, due to the limited extent of the identified data, and overall low quality, this was not performed.

2.2.7 Scoping of EarlyCDT Lung evidence outside the main diagnostic pathway

The database searches identified all published literature on the EarlyCDT Lung test. Given that the evidence identified was anticipated to be limited in both volume and relevance to the NHS setting, studies which did not formally meet the population inclusion criteria, or which fell outside the proposed diagnostic pathway (for example, where EarlyCDT Lung was used as a screening test) were deemed to be suitable for inclusion as part of a broader review, providing an eligible outcome was

reported. This additional literature is summarised narratively, where this literature informs understanding of the clinical impact of EarlyCDT Lung, or informs the economic analysis.

2.3 Cost-effectiveness reviews

2.3.1 EarlyCDT Lung for the diagnosis of lung cancer

The objective of this component of work was to perform a systematic review of published cost-effectiveness studies of EarlyCDT Lung for the diagnosis of lung cancer in patients with solid pulmonary nodules. Given a dearth of evidence on the cost-effectiveness of EarlyCDT Lung for lung cancer risk classification in patients with solid pulmonary nodules was expected, the review focussed on i) assessing the generalisability of available evidence to the decision problem defined by the NICE DAR scope and any particular positioning of EarlyCDT Lung in the diagnostic pathway, ii) identifying key structural and parameter assumptions, iii) identifying components of value of the technology and iv) characterising the evidence linkage mechanisms used to link these to final outcomes, in the existing cost-effectiveness models.

2.3.1.1 Literature searching

The results of the searches carried out for the systematic review of clinical effectiveness (see Section 2.2.1) were used to identify any relevant studies of the cost-effectiveness of EarlyCDT Lung for the diagnosis of lung cancer in patients with solid pulmonary nodules.

2.3.1.2 Study selection

A broad range of studies evaluating the cost-effectiveness of EarlyCDT Lung in the diagnostic pathway of lung cancer were considered, including economic evaluations conducted alongside trials, modelling studies and analyses of administrative databases. Only full economic evaluations that compare two or more options and consider both costs and consequences (including cost-effectiveness, cost-utility and cost-benefit analyses) will be included. Cost-effectiveness studies on EarlyCDT Lung used upstream from the diagnostic pathway of lung cancer (i.e., in screening for lung cancer) will be excluded from the review.

Studies identified by the search strategies (see Appendix 11.1) were screened in two stages. First, two reviewers (AD, MC) independently assessed and screened all records identified by the bibliographic searches for possible inclusion based on title and abstract, Second, full texts of potentially relevant studies publications were obtained for assessment and screened by two researchers (AD, MSo), with any disagreement resolved by consensus.

2.3.1.3 Quality appraisal

Cost-effectiveness evidence identified by the search was appraised for quality using a checklist specific to model-based economic evaluations of diagnostic tests.[16]

2.3.1.4 Synthesis of evidence

The characteristics and main findings of existing economic evaluations were narratively summarised and tabulated for comparison. In particular, information was extracted on:

- The comparators and positioning in the diagnostic pathway, study population, main analytic approaches (e.g. patient-level analysis/decision-analytic modelling) and primary outcome specified for the economic analysis;
- Key structural and parameter assumptions;
- Components of value (i.e., the features of the test in regards to comparators that allow establishing and quantifying trade-offs, the balance of which determines the net value of the technology);
- Details of adjustment for health-related quality-of-life (HRQoL), categories of direct costs and indirect costs;
- Estimates of incremental cost-effectiveness and approaches to quantifying decision uncertainty (e.g. deterministic/probabilistic sensitivity analysis).

The studies were critiqued in terms of their appropriateness and generalisability to inform the relevant decision problem (as defined by the NICE DAR scope), and whether they are particularly relevant for any of the proposed positionings for EarlyCDT Lung in the diagnostic pathway. The evidence linkage mechanisms used to link components of value to final outcomes will also be characterised, as part of the critique.

2.3.2 Additional targeted reviews to support model conceptualisation

To allow a fuller critical appraisal of the assumptions and data sources used in the existing cost-effectiveness studies and to assist in the conceptualisation of a new decision model, further targeted literature searches for cost-effectiveness studies were undertaken to identify a broader set of approaches (including relevant sources of evidence) for the evidence-linkage. These aimed to identify cost-effectiveness models evaluating other diagnostic strategies for lung cancer (such as those relating to the use of the Brock and Herder models or of PET-CT scan), and cost-effectiveness studies on screening approaches for lung cancer.

While this study's protocol stated that this review would restrict the inclusion of screening studies to those UK based studies, scoping reviews showed that this restriction might not provide sufficient

diversity of modelling approaches. Due to the high volume of literature in this area a pragmatic approach to developing the search strategy was taken, to ensure that the strategy was as inclusive as possible without retrieving an unmanageable amount of records for screening. The initial strategy was developed in Ovid MEDLINE combining terms for lung cancer screening or pulmonary nodules with a narrow study design search filter designed by CADTH to identify economic evaluations.[17] Text word searches of titles and abstracts were included in the strategy along with subject headings, some of which had focusing applied to increase the precision of the search. The MEDLINE strategy was translated to run appropriately on the other databases.

The following databases were searched on 24th March 2021: MEDLINE ALL (via Ovid - includes Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE Daily and Ovid MEDLINE), Embase (via Ovid), NHS Economic Evaluations Database (via CRD databases), Health Technology Assessment database (via CRD databases), International Health Technology Assessment database (via INAHTA website), and EconLit (via Ovid).

Records were imported into EndNote 20 for deduplication. All search strategies can be found in Appendix 11.6. Records were screened jointly for inclusion in either the diagnostic or the screening studies reviews. Screening was undertaken by a single reviewer (AD) in two stages on the basis of: i) title and/or abstract, and ii) full text publication. Publications were included in the diagnostics and screening reviews if they described studies evaluating the cost-effectiveness of alternative diagnostic and screening strategies for lung cancer, respectively, and met the following inclusion criteria:

1. Quantified cost-effectiveness using a decision analytic model;
2. Population of patients with SPNs and no previously diagnosed lung cancer;
3. Lung was the location of the primary cancer;
4. Used a linked evidence approach to quantify the impact of tests/screening on patient outcomes;
5. Publication written in English.

Furthermore, publications were excluded if these:

1. Consisted of conference abstracts, comments, editorials, notes, letters, errata or corrections;
2. Consisted of review articles without a *de novo* model or updated model;
3. Reported on the adaptation of existing models to deal with other disease populations (e.g., screening for lung cancer in human immunodeficiency virus positive patients).

Studies identified in these targeted reviews (both of diagnostic and screening models in lung cancer) were not subject to a formal assessment.

The evidence was synthesised narratively. In contrast with the review on the cost-effectiveness of EarlyCDT Lung (Section 2.3.1) study results were not summarised. The review focused on identifying value components relating to classification and describing the assumptions and data sources underpinning the linked-evidence approach, particularly those on the modelling of long-term health outcomes and costs. Key areas of uncertainty, evidential challenges, and UK relevant data sources were highlighted.

2.4 *Conceptualisation of the decision model and identification of evidence requirements for future assessments*

This component of work will focus on the conceptualisation of a decision model, structured according to good practice recommendations, [18, 19] to quantify the broader consequences to health and overall NHS and personal social services (PSS) costs associated with the use of EarlyCDT Lung (i.e., its cost-effectiveness). The recommendations and model conceptualisation will comply with the NICE reference case.[20] The key outputs of this element of work will be:

- outline key considerations for the development of an appropriate model structure, considering key structural assumptions and identifying the nature of the evidence linkages required, and
- an outline of key parameter inputs required, including an assessment of the possible data gaps that would need to be addressed in future research.

The conceptualisation process combined problem-oriented and design-oriented elements identified in Kaltehthaler et al., 2011.[19] The problem-oriented element of the conceptual modelling will describe: (i) current clinical understanding of the clinical condition and important events; and (ii) clinical pathways through which patients are detected, diagnosed, treated and followed-up. The design-led element of conceptual modelling will identify potentially feasible and credible model choices to represent the events and pathways deemed relevant in the problem-oriented element, considering the availability of existing evidence.

Explicit processes were used for the conceptualisation process, including interviews with a clinical expert and supported by the learnings from the suit of reviews conducted within this project. The results of the conceptualisation were recorded using influence diagrams[21, 22], which are reported in Section 6. Influence diagrams are compact representations of decision problems focussing on illustrating relationships between parameters in a model. These can be parameterised and implemented as decision analytic models (because of the probability-based representation of influence diagrams, these are typically translated into decision trees). However, we here use the influence diagrams to, more generally, reflect on relationships that need to be considered in a future assessment

of EarlyCDT Lung. These diagrams, therefore, are not to be used to convert the problem conceptualisation into an appropriate model structure, but to support further attempts in doing so as further evidence emerges updating knowledge of the disease, the technology and the process to be modelled.

The technology of interest is diagnostic, presenting a value proposition that is complex including indirect effects from changes in management decisions. The conceptualisation process was therefore structured to first identify value drivers and value components that could be of relevance for establishing the cost-effectiveness of EarlyCDT Lung in the diagnostic pathway for solid pulmonary nodules. Value drivers are here defined as factors, such as disease prevalence, that have expected to have a significant impact on cost-effectiveness. Value components are here defined as different mechanisms for clinical and economic impact of this technology (including any potential consequences of suboptimal treatment decisions in those misclassified) in this decision problem. These impacts may include,[23] for example, direct effects of the technology, effects derived indirectly by altering clinical decision on further tests or treatments, effects on the timing of decisions and actions, or influence on patient and clinician perspectives. These will include implications for resource use and for processes of health care service provision of the use of the test in relation to its alternative(s).

The conceptualisation then focussed on identifying possible mechanisms for evidence linkage for each of the components of value identified, for example, reflecting the consequences of diagnostic test accuracy as final cost and health outcomes.

3 DIAGNOSTIC ACCURACY AND CLINICAL EFFECTIVENESS RESULTS

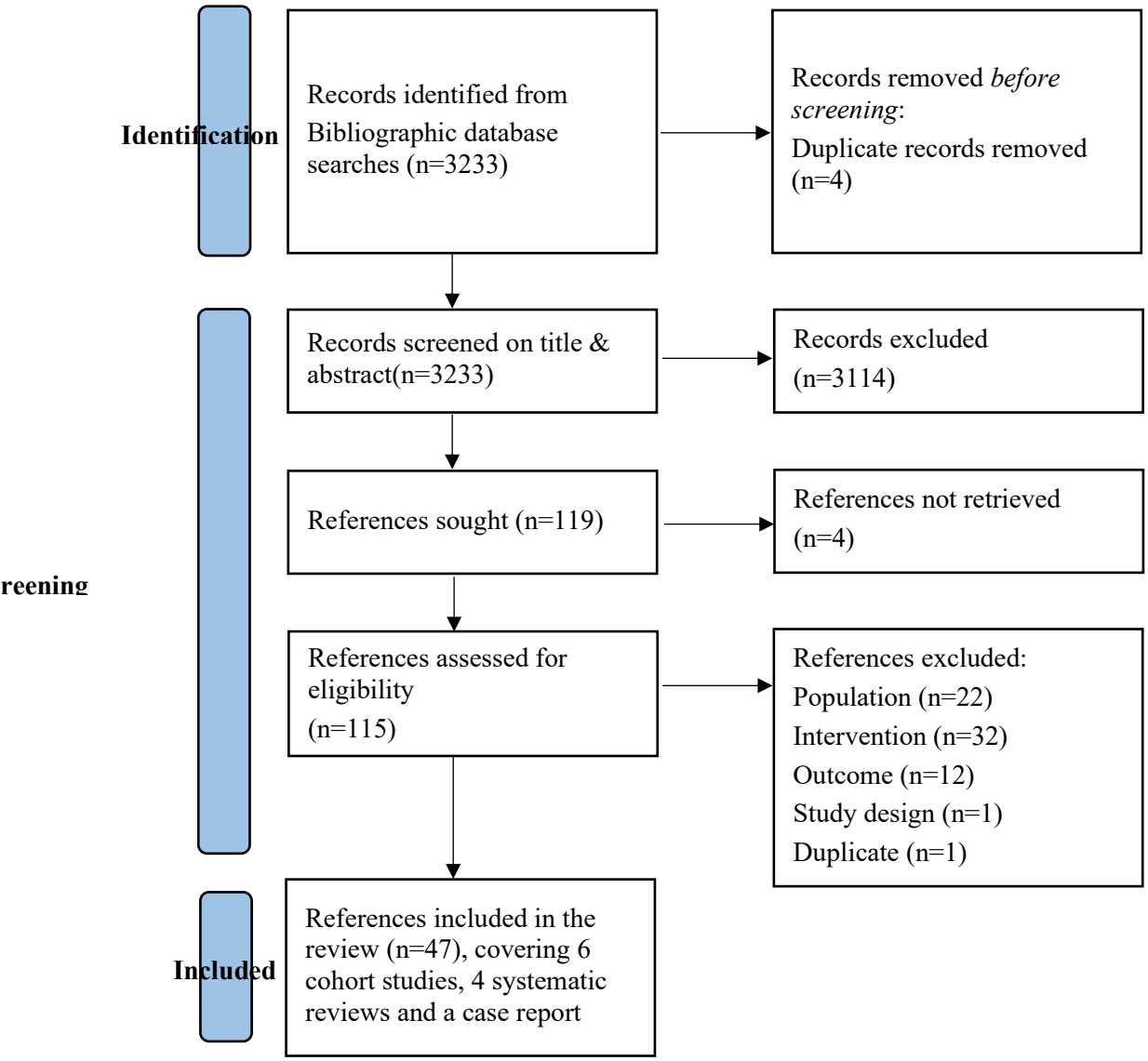
3.1 EarlyCDT Lung studies

3.1.1 Quantity of research available

Figure 4 presents a summary of the EarlyCDT Lung study identification and selection process. The searches identified a total of 3,233 unique records. After title and abstract screening, 115 references were retrieved, and 47 references were included in the review. Over half the included references were reported as conference abstracts.

Many references were excluded because the study populations consisted of patients with already diagnosed lung cancer (i.e. they studied validation cohorts of patients who would not receive the EarlyCDT Lung test in practice); seven published papers formed part of this group of references.[24-30] Although 47 references were identified as being eligible for inclusion in the review, they covered only six distinct patient cohorts, with some references reporting on subgroups within a cohort. See Appendix Table 32 for full details.

Figure 4 PRISMA flow diagram for the systematic review of EarlyCDT Lung studies



3.1.2 Cohorts using EarlyCDT Lung

Table 4 summarises the six patient cohorts and associated references for which diagnostic accuracy data were reported. The cohort associated with the most references was the study of 1,987 North American patients at high risk of developing lung cancer with Health Insurance Portability and Accountability Act (HIPAA) authorization, of which 451 had a pulmonary nodule. Results relating to this cohort were reported across five published papers between 2012 and 2017,[10, 31-34] and 10 conference abstracts between 2011 and 2018.[12, 35-43] A commentary article[44] was also published about one of the HIPAA cohort papers.[33]

Another cohort related to the EarlyCDT Lung Cancer Screening study of U.S. patients at high risk of developing lung cancer. This study, which compared CT scan alone vs. both the Early CDT-Lung test and CT scan, was registered on the clinicaltrials.gov registry[45] and has interim results reported, though only as conference abstracts.[8, 46, 47] The clinicaltrials.gov record (NCT01700257) shows an actual completion date of December 2020 and states that 1361 patients were recruited. The record also states that there will be "*health-economic costs included in the final analysis of study data*". As the latest of these conference abstracts was published in 2017 the EAG requested the NICE team to ask the manufacturer about the when the study results would be fully published. To date, these results have not been published, nor have results been submitted to the EAG. It is unclear when the results for this cohort will be published in full.

Two separate cohorts came from randomised trials in screening populations. One was based on patients recruited to the Early Diagnosis of Lung Cancer Scotland (ECLS) trial which randomised 12208 participants at risk of developing lung cancer. This cohort was reported in five published papers,[13, 48-51] a PhD thesis,[52] and 10 conference abstracts.[53-62]. One of the abstracts[59] also had a published erratum.[63] The other randomised trial cohort was based on a subgroup of 136 patients with pulmonary nodules who were recruited to a screening trial in Germany. The cohort of patients with pulmonary nodules was reported in one published paper.[64]

The final two cohorts were both very small: one was a U.S. study of 25 patients with indeterminate risk nodules, reported only as a conference abstract.[65] The other was a study of 10 Hong Kong patients with lung nodules, which was also reported only as a conference abstract.[66]

In addition to the cohort studies the searches identified four systematic reviews[67-70]. The systematic reviews included only studies identified by our search and only one included a meta-analysis of EarlyCDT Lung data, including only the HIPAA cohort data, so they were not considered further. We also identified a case report,[71] and a trial registry record for an ongoing study in China which is aiming to recruit 1000 patients.[72]

Table 4 Details of EarlyCDT Lung studies (and broader cohorts) which reported diagnostic accuracy outcomes

Reference and sample size	Reference type	Population	Period of sample collection	6 or 7 panel test?	Results for patients with lung nodules?
Cohort: North America Health Insurance Portability and Accountability Act (HIPAA) audit					
Jett 2011 [39] N=1010	Conference abstract	Patients at high risk of lung cancer	May 2009 - NR	6	No

Chapman 2012[31] N=836 (7 panel)	Paper	Patients at high risk of lung cancer	Nov 2010 - Aug 2011 (7-panel)	6&7	No
Chapman 2012[35] N=863	Conference abstract	Patients at high risk of lung cancer	NR	7	No
Healey 2012[36] N=847	Conference abstract	High risk patients (risk factors)	NR	NR	No
Jett 2012[40] N=861 (7 panel)	Conference abstract	High risk (based on Spitz model)	NR	6&7	NR
Healey 2012[37] N=959	Conference abstract	NR	NR	7	NR
Kucera 2012[41] N=70	Conference abstract	Patients at high risk of lung cancer	NR	NR	No
Peek 2012[43] N=108	Conference abstract	Lung nodule patients who tested positive for EarlyCDT lung	NR	NR	Yes
Healey 2013[32] N=847	Paper	Patients at high risk of lung cancer	Nov 2010 - Feb 2012	7	No
Massion 2013[42] N=423	Conference abstract	Lung nodules detected prior to EarlyCDT lung	NR	6&7	Yes
Jett 2014[33] N=861 (7 panel)	Paper	Patients at high risk of lung cancer	NR	6&7	No
Healey 2015[12] N=279	Conference abstract	Patients with CT detected lung nodules	NR	NR	Yes
Healey 2017[10] N=861	Paper	Patients at high risk of lung cancer	NR	7	Yes
Massion 2017[34] N=166 (7 panel)	Paper	Patients with lung nodules	May 2009 - Dec 2012	6&7	Yes
Jett 2018[38] N=48	Conference abstract	Patients with indeterminate lung nodules (risk >30%)	NR	7	Yes
Cohort: EarlyCDT Lung Cancer Screening study (LCS) NCT01700257					
Jett 2017[46] N=1235	Conference abstract	Patients at high risk of lung cancer	May 2012 - June 2016	7	No
Phillips 2017[47] N=1235	Conference abstract	Patients at high risk of lung cancer	May 2012 - June 2016	7	No
Jett 2015[8] N=815	Conference abstract	Patients at high risk of lung cancer	May 2012 - Nov 2014	7	No
Cohort: Early Lung Cancer Detection (ECLS) study RCT					
Sullivan 2021[13] N=12,208	Paper	Patients at high risk of lung cancer	April 2013 - July 2016	7	No
Cohort: U.S. indeterminate risk study					
Lin 2016[65] N=25	Conference abstract	Patients with indeterminate risk lesions	2014-2016	7	Yes
Cohort: Hong Kong pilot study					

Lau 2017[66] N=10	Conference abstract	Patients followed up for lung nodules	March - May 2017	7	Yes
Cohort: German screening RCT					
Gonzalez et al 2021[64] N=136	Paper	Patients with suspicious lung nodules	NR	7	Yes

NR Not reported

3.1.3 Summary of EarlyCDT Lung cohorts

Of the six cohorts identified where EarlyCDT Lung has been used it is important to note that none are explicitly of patients within the relevant BTS diagnostic pathway (Figure 1), as none explicitly reported that patients received a CT scan where nodules were identified which was then followed by an EarlyCDT Lung test.

In the HIPAA audit cohort most patients receiving an EarlyCDT Lung test did not have pulmonary nodules, and for those that did, it is unclear whether the nodules were identified before or after EarlyCDT Lung was performed. The study based on the German screening RCT cohort used a retrospective case-control design, with Early CDT-lung being performed on stored blood samples collected before cancer diagnosis.[64] In the ECLS trial EarlyCDT Lung was used as a screening test, prior to identification of nodules.[13] For the three cohorts only available as conference abstracts, it was unclear where EarlyCDT Lung was used in the diagnostic pathway.

Given that none of the cohorts met the strict inclusion criteria, this report instead focussed on analysis of the five cohorts (two with published papers, three with only conference abstracts) that reported data on patients with pulmonary nodules identified by CT scans.

A summary of the five cohorts is given in Table 5. The total sample size was small, with 695 patients with pulmonary nodules, including 97 diagnosed cancer cases. Cohorts had similar age distributions, and smoking rates. Three cohorts had broadly similar numbers of men and women, whereas two included mostly men.

Table 5 Summary of the included EarlyCDT Lung cohorts

Cohort	Primary data source	Location	Test threshold	Reference Standard	Number with nodules	Diagnosed cancers	Mean age	% Male	% current smokers
HIPAA	Massion 2017	USA	Commercial single threshold	Biopsy/surgery or 6 months follow-up	166	35	66	49	42
EarlyCDT LCS	Jett 2017	USA	Unknown *	24 month follow-up	352	7	59	45	52
US (Lin et al)	Lin 2016	USA	Unknown *	Biopsy/surgery or over 24 months follow-up	31	4	63	45	61
Hong Kong	Lau 2017	Hong Kong	Unknown *	Unknown	10	5	51.5	90	40
German RCT	Gonzales 2021	Germany	Double threshold from Healey et al 2017	Biopsy/surgery or over 24 months follow-up	136	46	63	70	52

** Unknown, but likely to be same as HIPAA*

3.1.4 Quality assessment

Table 6 summarises the results of the QUADAS-2 assessments for the cohorts with pulmonary nodules with full published papers. The Massion et al paper on the USA HIPAA cohort was judged to be at high risk of bias both in terms of patient selection (which was done by clinician judgement) and flow and timing (many patients were excluded from the analyses).[34] The paper by Gonzalez et al was also at high risk of bias for the patient selection domain.[64]This study utilised frozen blood samples taken at the time of the CT scan, but blood samples were not taken from 17 patients who went on to develop lung cancer (so these patients were excluded).

For both studies there were serious concerns about the applicability of their results to NHS practice. These concerns included the position in the pathway where the test was used (both studies), the way the test was used and interpreted (both studies) and use of a sub-optimal reference standard; the Massion et al study [34] followed up patients for only six months whereas the BTS guidelines recommend follow up of patients with nodules for one or two years.[3]

Table 6 Quality assessment of diagnostic accuracy studies reported in full published papers

Study	Risk of bias				Applicability concerns		
	Patient selection	**EarlyCDT Lung**	**Reference standard**	**Flow & timing**	**Patient selection**	**EarlyCDT Lung**	**Reference standard**
Massion et al 2017 (HIPAA audit cohort)[31, 34]	High	Low	Low	High	High	High	High
Notes	*Selection by clinician judgement (on risk of developing lung cancer). No pre-specified eligibility criteria nor protocol. The test is objective with pre-specified cut-offs. Many enrolled patients were excluded from analyses: follow up data outside the 6-month range (n=55), lost to follow up (n with nodules unclear), or nodule size not recorded (n=75).*				*Patients not initially included based on nodules and test was conducted before CT scan (up to a 6-month gap between the EarlyCDT lung test and CT scan). Test not used in patients with PET-CT scan data. Positive results not split by moderate and high thresholds. Follow up for only 'up to six months' (too short to identify all false-negatives).*		
González et al 2021[64]	High	Low	Low	Low	High	High	Low
Notes	*Although controls were randomly selected non-lung cancer patients with suspicious nodules, no blood samples were taken at CT for 17 excluded patients who went on to develop lung cancer. Lung cancer diagnosed before EarlyCDT lung test was done.*				*Test not used in patients with PET-CT scan data. Tests based on frozen blood samples. Pre-test and post-test risks not used. Long follow used up to detect all lung cancers.*		

HIPAA Health Insurance Portability and Accountability Act

Given the limited information presented in the conference abstracts quality assessments of the other three cohorts was not possible. It should be assumed that all three cohorts are at unclear risk of bias in all domains.

The ECLS study was a randomised trial which focussed on reporting clinical outcomes, so QUADAS-2 was not the most appropriate quality assessment tool. Assessment using the Cochrane Risk of bias tool found the trial to have a low overall risk of bias for the primary clinical outcome (see Appendix 11.2). However, many of the participants did not have pulmonary nodules - the trial was conducted in a high-risk screening population with the test result dictating whether CT imaging was performed. The results therefore have limited applicability to the population most likely to receive the EarlyCDT Lung test in NHS practice.

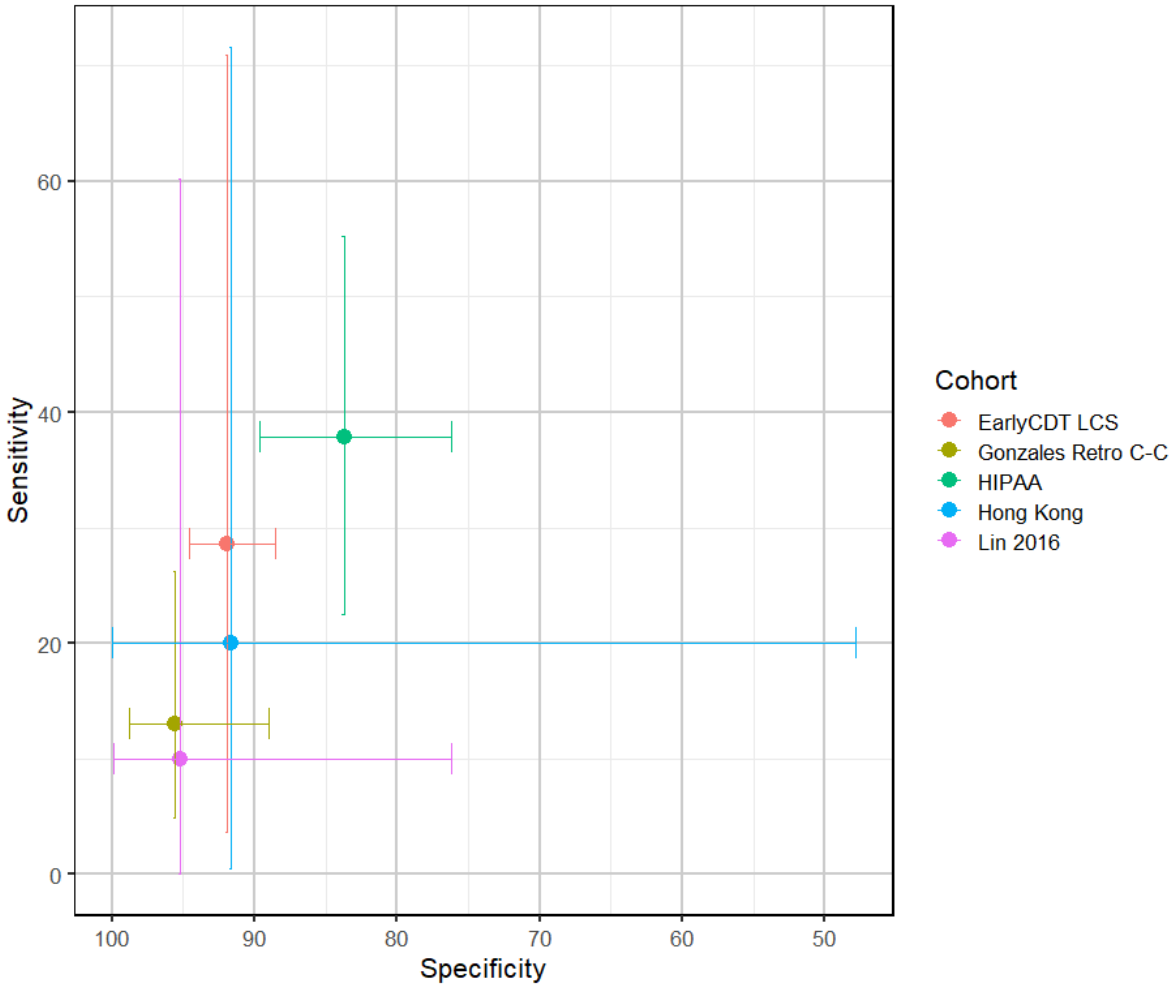
3.1.5 Synthesis of diagnostic accuracy

For 4 of the 5 identified cohorts of patients with pulmonary nodules diagnostic accuracy data were reported in one paper or abstract for each cohort. For the HIPAA cohort diagnostic accuracy data on patients with pulmonary nodules were reported in three papers. In this analysis we used data reported in Massion et al 2017 [34], as that was the most recently published and most comprehensive paper for that cohort.

The summary sensitivity and specificity data for the five cohorts are presented in Figure 5. The results of most cohorts are broadly consistent, with high specificity of over 90% but low sensitivity of under 30%. The HIPAA cohort showed higher sensitivity for lower specificity, but this may be because different test thresholds were used. The HIPAA cohort used the cut-offs for the commercial form of EarlyCDT Lung at that date. The Gonzales cohort used a “high-specificity” cut-off reported in another HIPAA paper (Healey et al 2017 [10]), this threshold is presented in Figure 5 and in the meta-analyses. Diagnostic accuracy data in patients with nodules were not reported for this “high specificity” cut-off in any HIPAA paper.

The Gonzales paper [64] also reported the diagnostic accuracy of using the combination of the “high specificity” and “moderate specificity” thresholds from Healey et al 2017 [10], which is the approach suggested by Oncimmune (see Figure 2). This found no change in sensitivity from using only the high specificity threshold (13%, 95% CI: 4.9 to 26.3), but a reduced specificity of 91.1 % (95% CI: 83.2 to 96.1) compared to 95.6% (95% CI: 89.0 to 98.8).

Figure 5 Diagnostic accuracy of EarlyCDT Lung from the 5 included cohorts



The summary sensitivity from univariate meta-analysis (see Figure 6) was 22% (95% CI 11 to 37). The summary specificity (see Figure 7) was 92% (95% CI 86 to 95). It should be noted that these estimates are based on different EarlyCDT Lung test cut-offs, and so may not represent test accuracy at any specified cut-off. As the test cut-off used was unclear for the three cohorts reported only as conference abstracts, a meta-analysis at specific test cut-offs was not possible.

The summary PPV was 32% (95% CI: 11 to 64). The summary NPV was 85% (95% CI: 63 to 95).). It should be noted that these summary results do not adjust for possible variation in prevalence across studies. The summary diagnostic odds ratio (DOR) was 3.32 (95% CI: 1.75 to 6.31). No study reported data on area under the ROC curve (AUC).

Figure 6 EarlyCDT Lung - meta-analysis of sensitivity

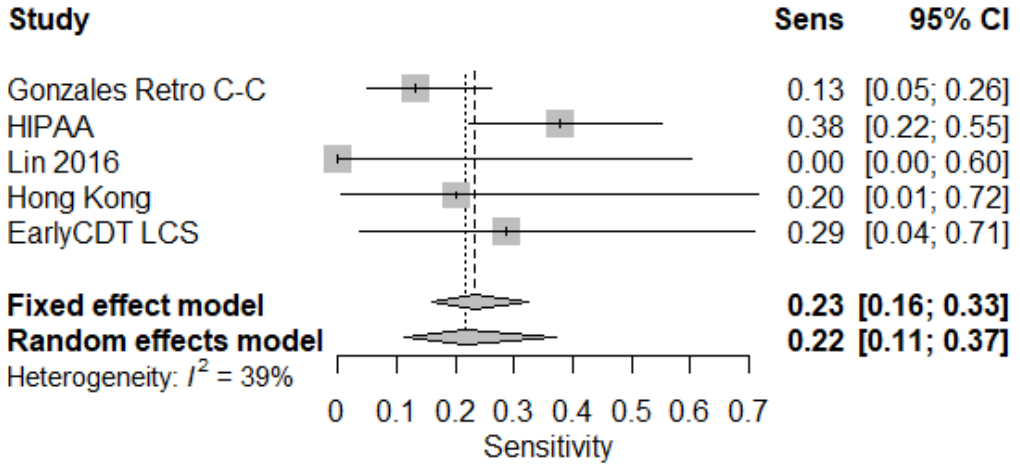
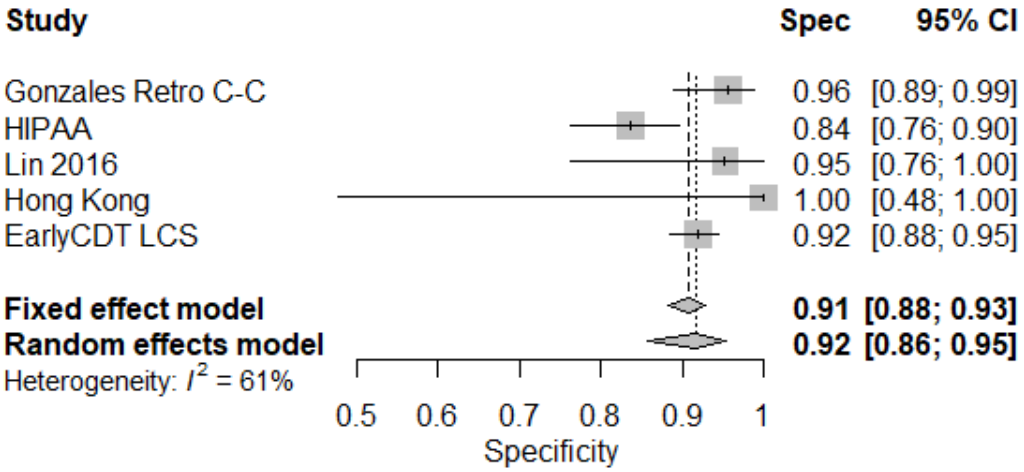


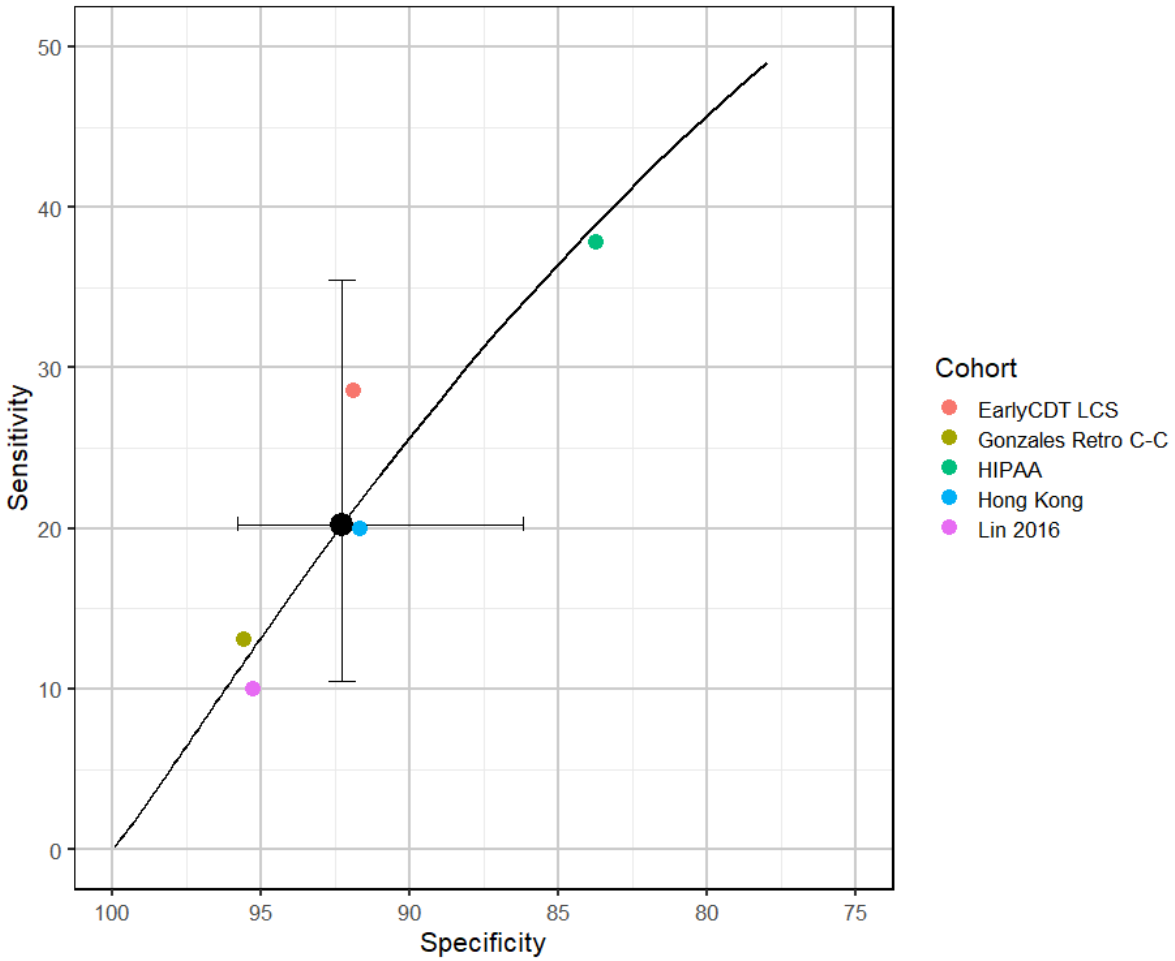
Figure 7 EarlyCDT Lung - meta-analysis of specificity



The results of a full bivariate meta-analysis of the EarlyCDT Lung cohorts, including a summary HSROC curve, are shown in Figure 8. The summary sensitivity in the bivariate model was 20.2% (95% CI 10.5 to 35.5) and the specificity was 92.2% (95% CI 86.2 to 95.8). However, as this includes cohorts using different EarlyCDT Lung cut-offs this may not be a reliable summary. Instead, from the HSROC curve we predict that Early CDT Lung has around 26% sensitivity at 90% specificity, or 12% sensitivity at 95% specificity. The area under the HSROC curve was 0.694, suggesting poor to moderate overall diagnostic accuracy.

The EAG notes that the diagnostic accuracy of EarlyCDT Lung in people with pulmonary nodules therefore appears to be poor and is lower than that predicted by Oncimmune (see Section 3.1.7.3). This may be because the risk models developed by Oncimmune were based on case-control studies of patients without pulmonary nodules.

Figure 8 EarlyCDT Lung - bivariate meta-analysis and HSROC curve



3.1.5.1 Diagnostic accuracy by nodule size

Diagnostic accuracy of EarlyCDT Lung by nodule size was reported only for the HIPAA cohort, here we consider results from the Massion et al 2017 paper, presented in Table 7. [34] Results were also presented in Healey et al 2017 [10] (see Table 9).

These results show no clear evidence of variation in diagnostic accuracy by nodule size, although it is possible that sensitivity declines, but specificity increases, with increasing nodule size. There were no malignant nodules with diameters below 4mm, so sensitivity could not be estimated for these smallest nodules. It does suggest that nodules are rare in this group so most positive EarlyCDT Lung results will be false positives (27.8% false-positive rate).

Table 7 Diagnostic accuracy by nodule size in Massion et al 2017

Nodule diameter	Sensitivity (95% CI)	Specificity (95% CI)
< 4 mm	(No malignant nodules)	72.2% (46.5 to 90.3)
4mm to 20mm	40% (16.3 to 67.7)	83.9% (74.4 to 90.9)
> 20mm	36.4% (17.2 to 59.2)	91.7% (73.0 to 98.8)

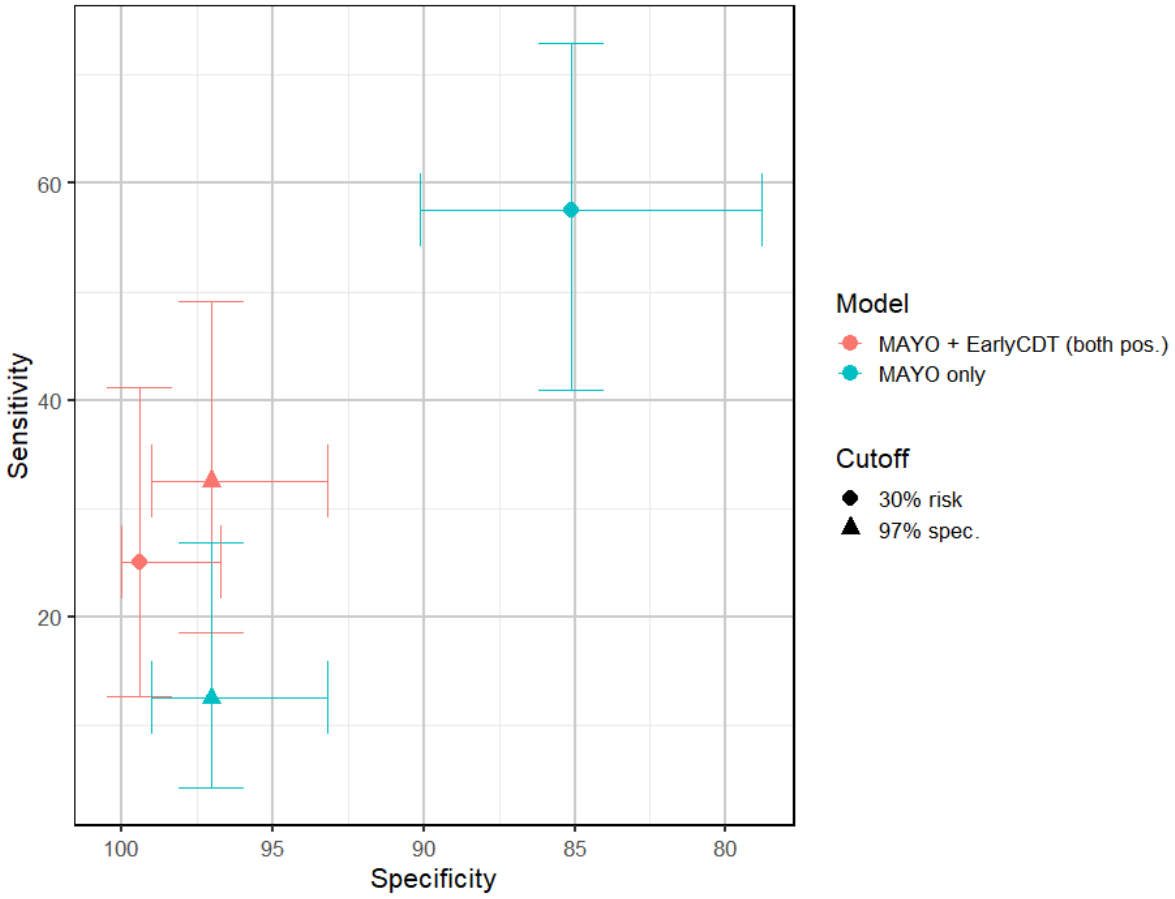
3.1.5.2 Combining EarlyCDT Lung with other risk scores

No studies reported any diagnostic accuracy data for the combination of EarlyCDT Lung with either the Brock or Herder risk assessment tools. Massion et al 2017 [34], reported data form the HIPAA cohort when combining EarlyCDT with the Mayo risk tool. This compared the Mayo risk alone to both Mayo and EarlyCDT being test-positive, at both 30% Mayo risk and an overall 97% specificity. The results are presented in Figure 9.

At 30% risk adding EarlyCDT Lung to Mayo substantially increased the specificity, but also decreased the sensitivity. At 97% specificity there is evidence that adding EarlyCDT Lung to Mayo risk can increase sensitivity. The paper does not state what risk level a specificity of 97% will equate to. Given that the specificity is much higher than at 30% risk it is likely to correspond to a high risk of malignancy.

It is not clear whether these results from using Mayo risk would be similar if Brock or Herder risk were used. Also, the “both positive” approach analysed here is not what is currently proposed for EarlyCDT Lung, where risk will be recalculated if EarlyCDT Lung is positive (see Figure 2).

Figure 9 Diagnostic accuracy of combining EarlyCDT Lung with Mayo risk



3.1.6 Synthesis of other clinical effectiveness outcomes

None of the five cohorts that reported on patients with pulmonary nodules presented any data on any of the broader clinical effectiveness outcomes (beyond diagnostic accuracy) listed in Section 1.6.

The ECLS screening trial reported that screening using EarlyCDT Lung resulted in earlier detection of malignant tumours than no screening [13]. However, as that was a screening study of people without identified nodules it is not possible to infer whether this earlier detection would also occur when assessing identified pulmonary nodules within the recommended BTS pathway. (See Section 3.1.7.1 for further discussion of the screening trial).

Therefore, the EAG concludes that there is currently no direct evidence on the clinical value of EarlyCDT Lung when used to assess pulmonary nodules.

3.1.7 Evidence on EarlyCDT Lung outside the diagnostic pathway

3.1.7.1 Early Diagnosis of Lung Cancer Scotland (ECLS) trial

The Early Diagnosis of Lung Cancer Scotland (ECLS) study was a randomised trial which addressed the question: "*Does using the EarlyCDT-Lung test to identify those at high risk of lung cancer and any subsequent CT scanning reduce the incidence of patients with late-stage lung cancer (III and IV) or unclassified presentation at diagnosis, compared with standard clinical practice*?"[13] This trial was undertaken in 12,208 individuals at increased risk of lung cancer (based on smoking history), with the intervention arm receiving the EarlyCDT Lung test and, if test-positive, low-dose six-monthly CT scans for up to two years. EarlyCDT Lung test-negative and control arm participants received standard clinical care (symptomatic presentation). The trial was therefore not designed to assess the incremental contribution of the EarlyCDT Lung test and although it did report sensitivity and specificity as outcomes the lack of a focus on a population, or subgroup, with pulmonary nodules means the diagnostic accuracy results are of very limited applicability to this assessment.

The ECLS study results have been reported across many conference abstracts and published papers (summarised in Table 8). The main trial paper by Sullivan et al[13] reported that 127 lung cancers were detected in the study population (1.0%) at 2 years. For the trial's primary outcome, in the EarlyCDT Lung test arm 33/56 (58.9%) lung cancers were diagnosed at stage III/IV compared with 52/71 (73.2%) in the control arm (hazard ratio for stage III/IV presentation: 0.64, 95% CI: 0.41 to 0.99). There were no statistically significant differences between groups in lung cancer mortality and all-cause mortality after two years. Five intervention-related adverse events (related to blood sample collection) were reported and all were considered to be minor.

Another ECLS paper looked at psychological outcomes in 338 patients who tested EarlyCDT Lung positive.[48] Responses of patients with pulmonary nodules on their first CT scan were compared to those without at three and six months. The paper reported no statistically significant differences between the groups in affect, lung cancer worry, health anxiety, illness perceptions, lung cancer risk perception or intrusive thoughts. Two papers[49, 51] reported on smoking behaviour outcomes following lung cancer screening.

Table 8 Summary of references reporting on the Early Diagnosis of Lung Cancer Scotland (ECLS) trial

Reference and sample size	Reference type	Population or subgroup reported	Period of recruitment	Outcomes reported
Sullivan 2021[13] N=12,208	Paper	Screening: Adults age 50–75 years at increased risk of developing lung cancer	April 2013 to July 2016	Sensitivity, specificity, mortality, adverse events, anxiety, depression, worry outcomes. Uptake of subsequent

				investigations such as CT or bronchoscopy.
Bedford 2017[52] N=1096	PhD Thesis	All EarlyCDT positives & random selection of negatives and control participants	Jan 2014 to May 2016	Positive and Negative Affect Schedule (PANAS), Lung Cancer Worry Scale (LCWS), Impact of Events Scale (IES)
Clark 2017[53] N=1032	Conference abstract	Sample of EarlyCDT positives, negatives and control participants	NR	Positive and Negative Affect Scale (PANAS), Lung Cancer Worry Scale (LCWS), Impact of Events Scale (IES)
Dorward 2016[54] N=12,018	Conference abstract	Adults aged 50-75 years who were at high risk for lung cancer	Completed in June 2016	EarlyCDT Lung test results and number of cancers
Sullivan 2015[56] N=12,000	Conference abstract	Adults aged 50-75 years at high risk of developing lung cancer	NR	EarlyCDT Lung test results and number of cancers. Current positive predictive value of test
Sullivan 2014[55] N=10,000	Conference abstract	Adults aged 50-75 years at high risk of developing lung cancer	NR	EarlyCDT Lung test results and number of cancers
Sullivan 2017[57] N=12,208	Conference abstract	Adults aged 50-75 years at high risk of developing lung cancer	Completed by June 2016	EarlyCDT Lung test results and number of cancers
Sullivan 2017[58] N=12,210	Conference abstract	Adults aged 50-75 years at high risk of developing lung cancer	NR	EarlyCDT Lung test results and number of cancers
Sullivan 2017[50] N=N/A	Paper (Protocol)	Adults aged 50-75 years at high risk of developing lung cancer	N/A	Protocol (list of outcomes)
Young 2017[60] N=1032	Conference abstract	Subsamples of EarlyCDT positives, negatives and control participants	NR	Smoking point prevalence, attempts to quit, number of cigarettes smoked per day and the Heaviness of Smoking Index
Young 2017[61] N=31 interviews	Conference abstract	Sample of people with positive and negative screening test results – either successfully or unsuccessfully attempted to stop smoking or no attempt since screening	NR	Qualitative interviews on facilitators to smoking cessation/cessation support – facilitators included emotional responses to test results
Young 2017[62] N=31	Conference abstract	Aged 51-74 years screened with the EarlyCDT Lung test (13 positive, 18 negative) and long-term smokers at screening	NR	Qualitative interviews: looking at how screening affected decisions about smoking, including interpretation of test results and emotional responses to results
Clark 2018[48] N=338	Paper	Subsample of EarlyCDT Lung positive participants (split between presence of nodules on first CT scan vs those without)	December 2013 to April 2015	Positive and Negative Affect Schedule (PANAS), Lung Cancer Worry Scale (LCWS), Health Anxiety Subscale (HAS) of the health orientation scale. Impact of Events Scale (IES). Revised illness perception questionnaire-adapted for lung

				cancer (IPQR) and lung cancer risk perception
Young 2018[51] N=31	Paper	Subsample of EarlyCDT Lung test participants (13 positive, 18 negative)	NR	Qualitative interviews: looking at how screening affected decisions about smoking, including interpretation of test results and emotional responses to results
Clark 2019[49] N=338	Paper	Subsample of EarlyCDT Lung positive participants (split between presence of nodules vs those without)	December 2013 to April 2015	Smoking behaviour
Sullivan 2019[59] N=12,210	Conference abstract	Adults aged 50-75 years at high risk of developing lung cancer		EarlyCDT Lung test results and number of cancers

NR Not reported, N/A Not applicable

3.1.7.2 Danish cohort of Borg et al 2021

Borg et al (2021) [73] performed EarlyCDT Lung on a cohort of 246 patients suspected of having lung cancer by their physician. This paper was published after our searches were completed. As patients did not have identified pulmonary nodules, and no data on patients with nodules was reported, the study is not eligible for the main analysis, and so is considered here.

All 246 patients received EarlyCDT Lung, with levels above either the "High" or "Moderate" thresholds described in Healey et al 2017 [10] being considered a positive result. Patients then had a CT scan and cancer diagnosis. All patients were followed up for a year to confirm or exclude cancer. The mean age was 65 years, with approximately equal numbers of men and women.; 76% of patients were current or former smokers. There were 75 diagnosed lung cancer cases (11 Stage I; 17 Stage II; 22 Stage III; 25 Stage IV).

The overall estimated diagnostic accuracy of EarlyCDT Lung was a sensitivity of 33% (95% CI 23 to 45) and specificity of 88% (95% CI 82 to 92). The paper reported diagnostic accuracy in several patient subgroups. The paper noted poor diagnostic accuracy for Stage I and II cancers (21% sensitivity for 88% specificity) and in patients aged 60 or under (11% sensitivity at 94% specificity).

The paper concluded that EarlyCDT Lung has insufficient sensitivity to be recommended as part of a low-dose CT lung cancer screening programme. The EAG notes that the paper does not report results for patients with pulmonary nodules, and inclusion was based on physician suspicion of cancer alone, so the study is not directly applicable to diagnosing pulmonary nodules. However, the low diagnostic accuracy in the study is consistent with that seen in the meta-analysis in Section 3.1.5.

3.1.7.3 Case-control studies of EarlyCDT Lung in patients without confirmed pulmonary nodules

A series of case-control studies were performed to assess the potential diagnostic accuracy of EarlyCDT Lung. These were reported in 2011 in papers by Boyle et al [24] and Lam et al [27]. All these initial case-control studies were of a different panel of autoantibodies to the current version of EarlyCDT Lung, consisting of a different set of six autoantibodies rather than the current seven. Also, as these were case-control studies, EarlyCDT Lung was performed after cancer diagnosis, and not after identification of pulmonary nodules. For these reasons the EAG considers these studies to be ineligible for inclusion in our main synthesis.

One of the case-control groups (235 cases and 236 controls, from UK, USA, Ukraine, and Russia) was subsequently re-tested using the current 7-panel version of EarlyCDT Lung, using stored serum samples. The results of this re-analysis were reported in Chapman et al 2012 [35], alongside some analysis of the included HIPAA cohort. The diagnostic accuracy for this re-evaluated sample was 41% sensitivity (95% CI: 35 to 48) at a fixed 91% specificity. As this analysis was not in patients with diagnosed pulmonary nodules it was also not included in our main synthesis.

Case-control studies may have substantial risk of bias when assessing diagnostic accuracy. This is because the test is performed after lung cancer diagnosis rather than before, and it is uncertain whether the test results (i.e. the levels of autoantibodies) would change over time, altering accuracy. The patients with cancer are unlikely to be representative of patients who would be included in a prospectively recruited cohort. The case-control study may be missing early-stage tumours which may be harder to diagnose with EarlyCDT Lung. Similarly, the control sample may not represent typical patients with benign nodules, particularly as patients were not matched on nodules characteristics, and control patients may not have had pulmonary nodules at all. The EAG therefore considers the case-control studies to be at high risk of bias for assessing diagnostic accuracy.

This risk of bias is particularly concerning as the case-control group assessed using the 7-panel version of EarlyCDT Lung were analysed again as part of the paper by Healey et al in 2017.[10] In that paper the case-control group (called the “optimisation cohort”) was re-analysed alongside data from HIPAA, including the subset of patients with pulmonary nodules which were included in our main synthesis. Diagnostic accuracy results were presented and are summarised here in Table 9 (based on Table 1 of Healey et al 2017).[10]

These results show that diagnostic accuracy from the case-control group was similar to accuracy in the overall HIPAA cohort. However, diagnostic accuracy in patients with pulmonary nodules was

notably worse than for the case-control group for all of sensitivity, specificity and likelihood ratio. This appears to be driven mainly by poorer diagnostic accuracy among smaller nodules, which are both more common than larger nodules in the HIPAA cohort, and more likely to be absent from the case-control group (because a cancer has to be diagnosed to be included). The EAG therefore considers there seems to be reasonable evidence that the diagnostic accuracy estimates from the case-control group may overestimate accuracy in patients with nodules.

In the Healey et al paper [10] the diagnostic accuracy for nodule and "optimisation" groups were claimed to be similar because Fisher exact tests found no evidence of difference (e.g. Fisher exact test for specificity: p=0.28). However, the number of patients in the nodule group was small (111 patients) and so the EAG considers that lack of evidence of a difference cannot be equated with no difference. The EAG therefore considers that it may be inappropriate to assume that diagnostic accuracy in the case-control group applies to patients with nodules.

Table 9 Diagnostic accuracy as reported in Healey et al 2017

Group	Sensitivity	Specificity	Positive likelihood ratio
Case-control	41.3% (35.0 to 47.6)	90.6% (87.1 to 94.1)	4.4 (2.9 to 6.6)
HIPAA (all patients)	47.4% (24.9 to 69.8)	90.5% (88.4 to 92.5)	5.0 (3.0 to 8.3)
HIPAA (with nodules)	37.8% (22.2 to 53.5)	85.6% (79.1 to 92.1)	2.6 (1.4 to 4.8)
Small nodules (4-20mm)	40.0% (15.2 to 64.8)	83.9% (76.2 to 91.6)	2.5 (1.1 to 5.4)
Larger nodules (> 20mm)	36.4% (16.3 to 56.5)	91.7% (80.6 to 100)	4.4 (1.0 to 18.4)

The Healey et al risk model

The case-control group in Healey et al 2017 [10] was then used to construct two new sets of EarlyCDT Lung test thresholds: the "high specificity" threshold (98% specificity, with 28% sensitivity) and "low specificity" threshold (49% specificity for 80% sensitivity). The risk model proposed for general use (see Figure 2) was constructed assuming the stated diagnostic accuracy for these two new thresholds is valid. If it is, in fact, an overestimate of the diagnostic accuracy, then the post-test risk estimated by these models will be too high and decisions made using the rule may be invalid.

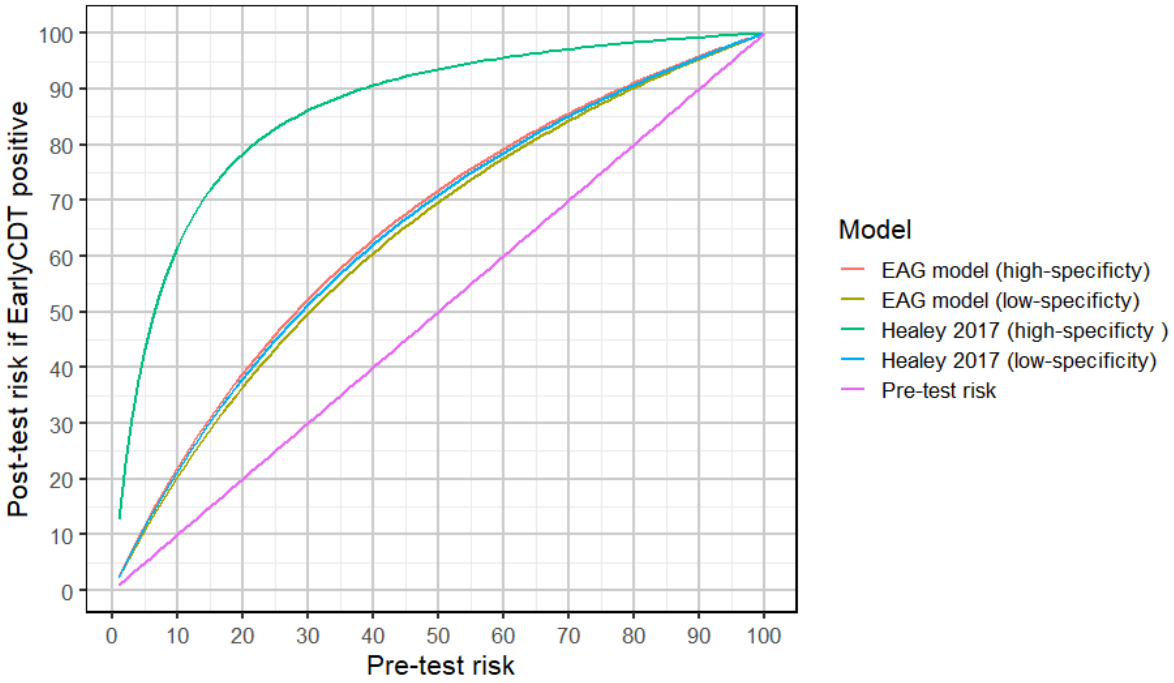
As the EAG meta-analysis does not support these estimates of diagnostic accuracy we compare the risk model form Healey et al to an "EAG model" with sensitivity estimates taken from the bivariate

meta-analysis (Figure 8) at the same specificity thresholds (Sensitivity 5.1% at 98% specificity; sensitivity 46% at 80% specificity). We note that this analysis does not account for uncertainty in diagnostic accuracy, in either the EAG analysis, or that of Healey et al.

For people with negative EarlyCDT Lung the post-test risk was assumed to be unchanged from the pre-test risk. For people with positive EarlyCDT Lung results the post-test risk was calculated from pre-test risk and diagnostic accuracy as set out in Healey 2017 [10]. Briefly, the pre-test risk and sensitivity/specificity of EarlyCDT Lung were combined to estimate the true positive and false-positive rates, and these used to calculate the positive predictive value, which was taken to be the post-test risk.

The pre-and post-test risks for the model using the "high specificity" and "low specificity" thresholds from Healey et al 2017 [10], and from the EAG model are presented in Figure 10. The increase in risk if EarlyCDT Lung is positive is much smaller for the EAG model, for the "high-specificity" threshold, because of the much lower predicted sensitivity. Consequently, a positive EarlyCDT Lung test is less likely to change a patient's risk classification in the EAG model (e.g. from low (<10%) to intermediate risk (10-70%)).

Figure 10 Post-test risk using the Healey 2017 and EAG models



3.2 Comparators

Database searches for all comparators described in Section 1.4 were performed. These searches produced 3,647 potentially relevant publications. Given the size of this database, and the limited evidence on EarlyCDT Lung, it was decided to perform targeted screening of these identified publications using keyword searches in Endnote to identify relevant papers.

Keyword searching was used to identify all likely systematic reviews or meta-analyses (91 papers), and these were screened for inclusion. For comparators where systematic reviews were not identified, further keyword searches were performed to identify relevant individual studies.

3.2.1 Small nodules

We identified no systematic reviews of patients explicitly with small nodules (5-8mm in diameter or 80-300mm^3 in volume).

We identified one study (Al-Ameri et al) [74] that reviewed the outcomes of 211 patients with pulmonary nodules undergoing diagnosis for lung cancer. The study reported that 37 of 211 patients had nodules 5-8mm in diameter who were all referred to CT surveillance. The number of malignant tumours in these patients was not reported. 6% of all patients in CT surveillance had malignant tumours, so it is unlikely that more than 6% (i.e. 2 of the 37) of patients with small nodules had malignancies.

Clinical advice to the EAG was that small nodules tend to be more difficult to biopsy, or may not be amenable to biopsy in some circumstances, so CT surveillance will be the normal management approach for such nodules. See Section 3.2.6 for discussion of CT surveillance.

3.2.2 The Brock risk model

We identified no systematic reviews or meta-analyses of the Brock risk assessment model. Targeted keyword searching for “Brock” or “PanCan”(an alternative name for the test) within our Endnote database of comparator studies identified 28 possibly eligible papers, of which 9 reported data on the diagnostic accuracy or clinical effectiveness of the Brock model. Studies in Asian populations were excluded as several showed evidence that the Brock model has inferior accuracy in East Asian countries, so were deemed not relevant to the UK context. [75] Other studies were excluded as no full text was available, no relevant accuracy data was reported, or for multiple publications of the same cohorts.

A summary of the 9 included publications is given in Table 10. Three of these papers reported data from the National Lung Cancer Screening Trial (NLST) cohort. It was unclear whether the papers

analysed the same, or different, patients within the wider cohort. For completeness, we report the results of all papers here.

Studies were a mix of prospective and retrospective cohorts, with two case-control studies and one clinical trial. All appeared to use a reasonable reference standard of biopsy or surgery or clinical follow up to confirm the presence or absence of cancer. In all studies the CT scan was performed before diagnosis, but in retrospective cohorts the Brock risk calculation will have been performed after diagnosis. This is unlikely to lead to substantial bias, given that Brock risk is based on results of the CT scan.

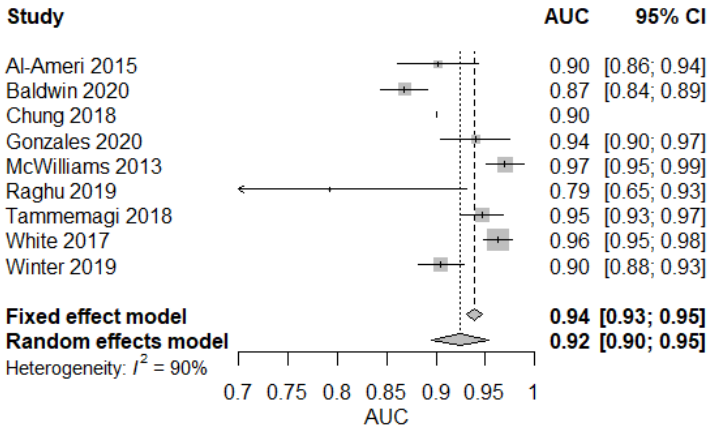
Table 10 Summary of papers reporting diagnostic accuracy data for the Brock model

Paper	Cohort	Location	Design	Size	No. cancers	Mean age	Percentage male	Percent recent or current smokers
Al-Ameri 2015	Independent	UK	Retrospective Cohort	244	99	69	50	76.2
Baldwin 2020	Independent	UK	Retrospective Cohort	1187	229	65 (approx)	51.3	NR
Chung 2018	Independent	Netherlands	Retrospective Case-control	1786	381	63 (approx)	47.4	NR
Gonzales 2020	LUSI (trial)	Germany	Clinical trial	1159	62	57.6	65.8	100 (61.9 current)
McWilliams 2013	PanCan / BCCA	Canada	Prospective Cohort	2961	144	NR	NR	NR
Raghu 2019	PLuSS	USA	Prospective nested case-control	50	42	64	57	55 (current)
Tammemagi 2018	PanCan / NLST	Canada / USA	Prospective Cohort	1711	111	62.5	53.1	NR
White 2017	NLST	USA	Prospective Cohort	2819	116	62	61	NR
Winter 2019	NLST	USA	Prospective Cohort	5018	194	63.7	61.3	NR

3.2.2.1 Meta-analysis

Studies generally reported the area under the ROC curve (AUC) to summarise diagnostic accuracy, rather than using sensitivity and specificity. This accounts for the fact that the Brock model might be assessed at different risk cut-offs (e.g. 5%, 10%). The forest plot for the meta-analysis of reported AUC values is given in Figure 11. This suggests that the Brock model has very good diagnostic accuracy (AUC 92%, 95% CI: 90% to 95%), but with some evidence of heterogeneity across studies (I^2 = 90%), with estimated AUCs varying from 79% to 96%. We note that AUC does not provide evidence on the diagnostic accuracy at specific cut-offs of interest (such as the 10% risk cut-off). A sensitivity meta-analysis excluding two of the three papers reporting data on the NLST cohort, and retaining only the most recent (Winter 2019 [76]) had a very similar result (AUC 91%, 95% CI: 87% to 95%)

Figure 11 Forest plot of AUC values for studies assessing the Brock risk model

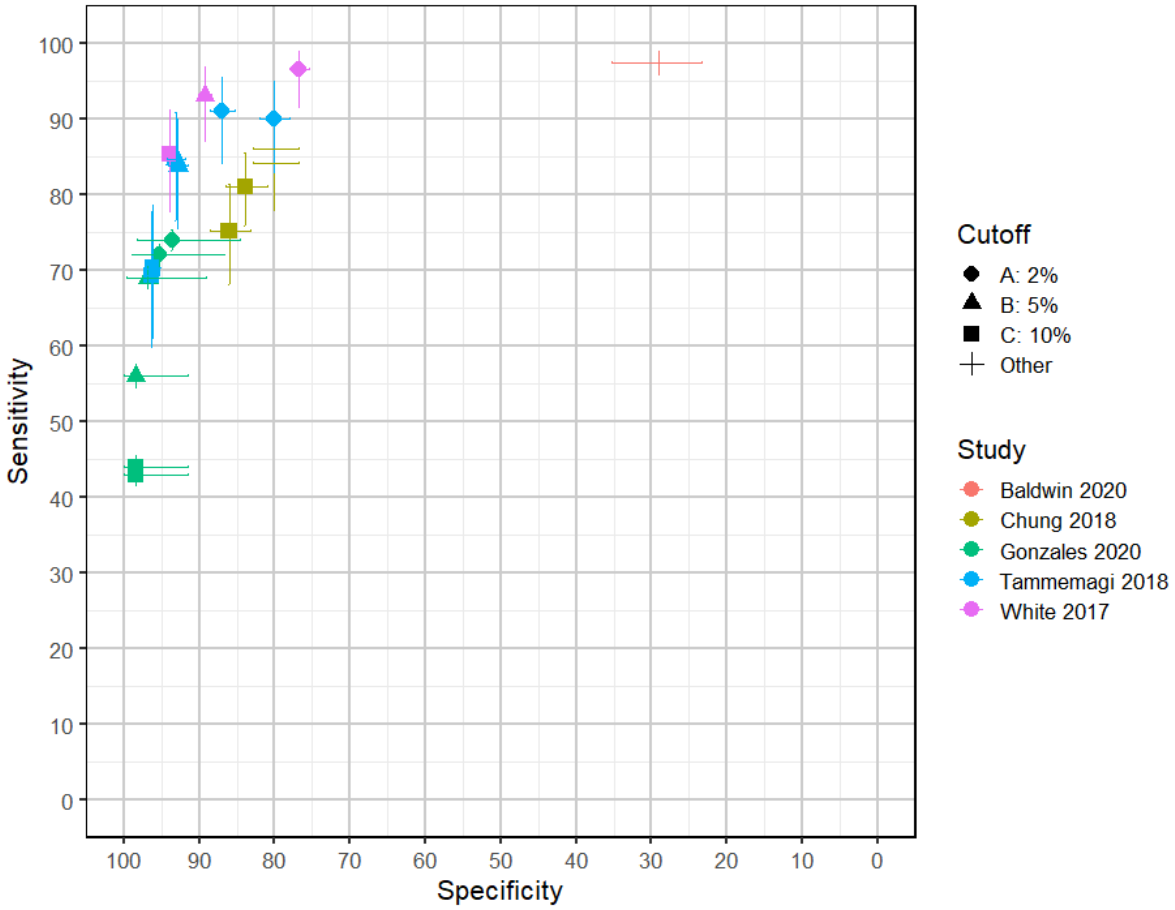


** The Chung 2018 study did not report a confidence interval for AUC, so was not included in the meta-analysis*

Five of the included studies reported sensitivity and specificity estimates for the Brock model at various thresholds. These are plotted in Figure 12. There was some heterogeneity across studies, even when using the same threshold of risk (e.g. 10%, the squares in Figure 12), but all studies suggest high diagnostic accuracy, with 80% sensitivity at 90% specificity appearing to be achievable. This contrasts with the estimated 25% sensitivity at 90% specificity for EarlyCDT Lung.

As there were only two cohorts (NLST and LUSI) reporting sensitivity and specificity at most risk thresholds, and as these had heterogeneous results (see Figure 12), no meta-analysis of sensitivity or specificity is presented here. Consequently the diagnostic accuracy of the Brock model at any particular risk cut-off (such as the 10% cut-off to distinguish low risk and intermediate risk nodules) is uncertain.

Figure 12 Sensitivity and specificity estimates by test threshold for the Brock model



3.2.3 The Herder risk model

We identified no systematic reviews or meta-analyses of the Herder risk assessment model. Targeted keyword searching for "Herder" or "Mayo" within our Endnote database of comparator studies identified 7 possibly eligible studies, of which 4 reported data on the diagnostic accuracy or clinical effectiveness of the Herder model explicitly. Given this limited number of studies we also included two studies that reported diagnostic accuracy when combining PET-CT with the Mayo risk tool,

which is functionally very similar to the Herder model. As for the Brock model, studies in Asian populations were excluded. One study was excluded as it reported no relevant data.

A summary of the 6 included cohorts is given in Table 11. All 6 papers reported data from different cohorts. All studies were retrospective cohort studies. All appeared to use a reasonable reference standard of biopsy and surgery or clinical follow up to confirm the presence or absence of cancer. In all studies the PET-CT scan was performed before diagnosis, but the Herder or Mayo risk calculation will have been performed after diagnosis. This is unlikely to lead to substantial bias, given that Herder risk is based on results of the PET-CT scan.

Table 11 Summary of papers reporting diagnostic accuracy data for the Herder risk model

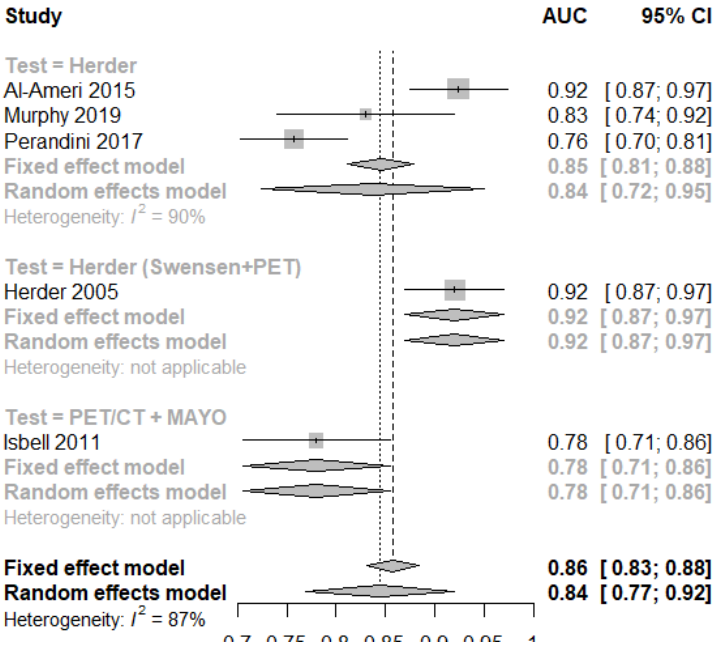
Paper	Test	Location	Size	No. cancers	Mean age	Percentage male	Percent recent/current smokers
Al-Ameri 2015	Herder	UK	244	99	69	50	76.2
Herder 2005	Herder	Netherlands	106	61	63	33	75
Murphy 2019	Herder	UK	97	75	69	52	84
Perandini 2017	Herder	Italy	259	153	66	64	90
Evangelista 2013	PET/CT + MAYO	Italy	59	31	70	54	54
Isbell 2011	PET/CT + MAYO	USA	189	138	63	50	74

3.2.3.1 Meta-analysis

As for the Brock model, most studies presented results as summary AUCs. A forest plot of these results is shown in Figure 13. These results suggest good diagnostic accuracy for the Herder model overall, with an AUC of 84% (95% CI 77% to 92%). There was substantial heterogeneity. Notable was the much lower accuracy seen in the Perandini study [77] than in earlier studies. That paper acknowledged this difference, but could not explain the heterogeneity.

Only 3 studies reported any sensitivity or specificity estimates for the Herder model. These are summarised in Appendix Figure 25. Again, these suggest only moderate to good diagnostic accuracy, of approximately 50-60% sensitivity at 90% specificity. Data were too limited to perform meta-analyse sat any specific risk cut-offs.

Figure 13 Forest plot of AUCs for the Herder risk model



3.2.4 Brock and Herder risk models

The Al-Ameri 2015 study was the only study that reported results for both Brock and Herder models. In the study 244 patients had a CT scan; 139 patients had a PET-CT scan provided nodule size was >5mm and physician judged that the risk of malignancy justified performing PET-CT (so not necessarily in line with BTS guidance).

The Brock model had an AUC of 0.902 (95% CI: 0.856 to 0.948, n = 154) and Herder had an AUC of 0.924 (95% CI: 0.875 to 0.974, n = 113). The Herder model AUC was statistically significantly superior to that for the Brock model (p = 0.002). At 90% specificity the sensitivity of the Brock model was approximately 70% (digitally extracted from ROC curves) and was approximately 78% for the Herder model.

3.2.5 Studies of the use of PET-CT in patients with pulmonary nodules

Targeted keyword searches of the Endnote database of comparators, other diagnostic tests in the care pathway and surveillance, together with citation searching in Google Scholar, identified 9 recent (i.e. published after 2010) systematic reviews[78-86] and one review of meta-analyses[87] of studies reporting diagnostic accuracy data for PET-CT in patients with pulmonary nodules.

A review of recent meta-analyses was published in 2020 by Lococo et al.[87] Two bibliographic databases were searched from 2010 onwards. The review included 10 meta-analyses of studies which reported results based on a single timepoint and four meta-analyses which looked at outcomes after

dual timepoints. All the single timepoint meta-analyses except one[78] reported sensitivities as being higher than specificities. The Lococo et al study also concluded that the data did not support the routine use of dual time point assessments. However, this study lacked details to inform an assessment of the quality of the meta-analyses. Given this, and the likely duplication of included studies across the different meta-analyses, a focus was made on the more recently published and largest meta-analyses: Jia et al 2019[80] and Li et al 2018.[81]

In the Jia et al review[80] three bibliographic databases were searched and two independent reviewers screened studies and extracted data. The review included 23 studies covering a total of 2024 patients. A quality assessment of the included studies was performed, though it utilised the out-of-date original QUADAS tool, rather than the updated QUADAS-2, which was published in 2011. QUADAS-2 enhanced the original tool substantially by utilising 'signalling questions' to improve both judgments and reporting transparency on key aspects of bias and applicability (which were considered as distinct aspects of study quality).[88] Nevertheless, using QUADAS, nearly all the included studies were deemed to be of moderate or high quality though it was not possible to verify this, given the limited details reported. Nine of the 23 included studies were conducted in China, Japan or Korea with most of the remainder conducted in the U.S. or Europe (one study was set in the UK). Sixteen studies had a sample size of less than 100; the largest study recruited 298 patients. Minimum nodule sizes in patients ranged between 1.4 and 10mm and maximum sizes were 30mm in nearly all studies. Standardized uptake value (SUV) thresholds were not reported.

In the Li et al review[81] two bibliographic databases were searched and two independent reviewers screened studies for eligibility; the data extraction methods used were not clearly reported. The review included 21 studies covering a total of 1557 patients. QUADAS-2 was reported as being used for quality assessment but it appeared that this was not the case; a modified version of the original QUADAS tool appeared to have been used with each study given a score (maximum of 11, it appeared that most studies scored ≥9). Six of the 21 included studies were conducted in China, Japan or Korea; one study was set in the UK (the same study which was included in Jia et al's review[80]). Seventeen studies had a sample size of less than 100; the largest study recruited 218 patients. Maximum standardized uptake value (SUV_{max}) cut-offs to determine malignancy status were 2.5 MBq/g in many studies, though ranged from 1.5 to 24 MBq/g.

Diagnostic accuracy results from both reviews are presented in Table 12. Sensitivities are higher than specificities in both meta-analyses and results are broadly similar across outcomes, although this would be expected, given the large overlap of included studies across the two reviews. It appeared that none of the studies in these meta-analyses reported the performance of PET-CT based on nodule

size or on pre-test likelihood of malignancy, as categorised in clinical guidelines. Citation searches were therefore conducted for any recent studies which reported such data.

Table 12 Results of two recent meta-analyses on the diagnostic accuracy of studies of PET/CT in patients with pulmonary nodules

Outcome	Meta-analysis results*	
	Jia et al 2019,[80] 23 studies, N=2024	**Li et al 2018,[81]** ** **21 studies, N=1557**
Sensitivity	0.89 (0.85 to 0.92), I^2 NR	0.88 (0.85 to 0.90) I^2=75%
Specificity	0.78 (0.66 to 0.86), I^2 NR	0.67 (0.63 to 0.71) I^2=89%
LR+	3.97 (2.57 to 6.13), I^2 NR	3.09 (95% CI: 2.11 to 4.52) I^2=90%
LR-	0.15 (0.10 to 0.20), I^2 NR	0.20 (95% CI: 0.13 to 0.29) I^2=72%
DOR	24.04 (12.71 to 45.48), I^2=79%	18.47 (95% CI: 8.75 to 38.97) I^2=81%

*95% confidence intervals are in brackets, DOR Diagnostic odds ratio, LR+ Positive likelihood ratio, LR- negative likelihood ratio, NR Not reported, ** Reported results differ between text and forest plots – forest plot results are reported here

3.2.5.1 Primary studies of PET-CT with results stratified by pre-test risk or nodule size

Two relevant and recent studies which stratified results by pre-test risk or nodule size were identified. Evangelista et al 2018 conducted a retrospective study of the diagnostic accuracy of FDG PET-CT stratified by risk of malignancy - based on the Brock model - in 502 Italian patients with solitary pulmonary nodules (identified by CT images) between 6mm and 30mm in diameter.[89, 90] Patients with indeterminate results (N=147) were excluded from the analyses, leaving a sample of 355 patients. Final diagnosis was made by histopathology (of biopsied or excised tissue) and/or by other imaging data at follow-up. Nodules that did not change or spontaneously resolve during 24 months' of follow-up were considered benign. Sensitivity and positive predictive value were significantly higher ($p<0.05$) in intermediate-risk (5-65%) and high-risk (>65%) patients, than in low-risk (<10%) patients while specificity and negative predictive value were significantly higher ($p<0.05$) in low-risk patients than in the other risk subgroups. Results for each risk group were also reported based on different methods of measuring standardized uptake value (SUV) ratios (mediastinal blood pool and liver).

In a prospective UK multicentre trial (called SPUTNIK) Weir-McCall et al 2021[91] compared qualitative and semi-quantitative PET/CT criteria, and the impact of nodule size on the diagnosis of solitary pulmonary nodules ≥ 8 and ≤ 30mm. The presence/absence of lung cancer was based on biopsy/histology or the completion of two years of follow-up. Sensitivity and specificity results were presented by optimised nodule size-specific (<12mm, 12-16mm and >16mm) cut-points for SUV_{max}, SUR_{blood} and PET grade. The study recruited 360 patients and concluded that SUV_{max} was the most accurate technique for the diagnosis of solitary pulmonary nodules and that diagnostic thresholds should be altered according to nodule size. However, the study was limited in its relevance to this

assessment in that patient pre-test probability of malignancy was not stratified based on the Brock model and the Herder model was not used to inform an estimate of post-test risk of malignancy.

3.2.6 CT surveillance

Targeted keyword searches in our EndNote database of comparators did not identify any systematic reviews or individual cohort studies specifically on using CT surveillance within the pulmonary nodule diagnostic pathway (e.g. in patients with low-risk nodules or nodules of small size).

The Al-Ameri study [74] stated that CT surveillance was recommended for 106 patients out of a total sample of 211. (37 with nodule <8 mm, 45 with malignant risk of <10% following initial CT, and 24 with malignant risk of <10% following PET-CT). Across all of these groups, six had malignant tumours within 2 years of follow-up, suggesting around a 6% risk of malignancy in those referred to CT surveillance. No other studies of CT surveillance within the framework of the BTS guidelines were identified.

Given the limited evidence on CT surveillance directly we performed targeted searches within our database for studies where CT was used as a screening tool for lung cancer (i.e. in patients at risk of cancer but with no identified nodules at time of screening), and reviewed all such studies that reported on the impact of further CT screening (e.g. after 3 months, 6 months or 1 year) in people with identified nodules. We identified five such studies. Given the variation in reporting of these studies, no meta-analysis was feasible, so we present a narrative summary of the studies.

3.2.6.1 Manchester Lung Health Check

This study has reported both its initial round (Crosbie et al 2018 [92]) and a second round (Crosbie et al 2019 [93]).

The numbers of people testing positive (i.e. requiring further testing such as PET-CT or biopsy) after a 3-month follow-up CT scan after their first CT scan was around 8.5% (16 of 189). The rate was similar for 3-month follow-up CT after the second screening round 1 year later (6 of 69) All malignant tumours identified at the 3-month follow-up CT scans were stage I tumours. Data were not reported for other follow-up CT scans after 3 months.

Some of these patients will be false-positive results (test positive but with benign nodules). There may also be false negatives. There was insufficient data in the published papers to determine the numbers of these. However, of all people referred to cancer clinics on the basis of any CT scan (including those referred after the first screening CT scan), 48.3% were false-positive results.

3.2.6.2 UK Lung Cancer Screening (UKLS)

Field et al [94] reported the key findings on patients with follow-up CT scans in the UKLS cohort. In this screening trial patients with small nodules (15-50mm^3 or 3-5mm) identified at first screening had a follow up CT at 1 year; patients with larger nodules (50-500mm^3 or 5-10mm) had a follow up CT at 3 months.

3.2.6.3 NELSON

Walter et al (2019) [95] presented results from the NELSON trial of CT screening, where volume doubling time and nodule size were used to diagnose cancer in patients with detected nodules undergoing CT at 3 months and approximately 1 year after initial screen. Histopathology was used to confirm malignancies. The reference standard for benign nodules was no detected tumour at any NELSON screening round, or in subsequent patient records. Diagnostic accuracy of this approach is summarised in Table 13 (after Table 3 of Walter et al [95]). The combination of volume and volume doubling time produces extremely high sensitivity to detect malignant nodules (estimated at 100%). The specificity is also high.

Of the 42 diagnosed cancers, 10 were diagnosed at the follow-up CT scan (24%), with the rest identified at first screening. It is unclear whether there were any false-negatives (cancers undetected after 2 years). 472 patients had a 3-month follow-up CT, 43 were referred for further tests, 9 (2%) had cancer (2 diagnosed immediately, 7 after 1-year follow-up); 479 patients had a 1-year follow-up, 7 were referred for further testing, 1 of whom had cancer. Six of seven the cancers identified after one year were at Stage I when detected, with one at Stage IV.

3.2.6.4 West London Screening pilot

Bartlett et al (2020) [96] reported preliminary results from this screening trial. Of 163 patients with an initially indeterminate CT scan, 143 have since undergone further CT or PET scans (at 6 weeks, 3, 6, or 12 months). Of these, 15 had a positive CT, and 10 to 15 (number not given) were subsequently diagnosed with cancer. 29 had results that were still indeterminate, and 102 were negative. Further testing has been delayed by COVID.

3.2.6.5 LUSI

Becker et al (2020) [97] reported aspects of the LUSI screening trial in Germany. After the first screening round 19.6% of screened persons were followed up with a CT scan at 3 or 6 months. The number of cancers among these persons was not stated, but was under 25 (<6%).

At subsequent annual screening rounds the recall rate (immediate, 3 or 6 months) ranged from 4.0 to 5.7%. The decline relative to the first round was attributed to the use of volume doubling time to

assess risk, which appears to have improved the specificity of screening. Around 13-15% of recalled patients were diagnosed with cancer (about 0.6% of the total screened, per year). Interval cancers, undetected by CT screening were rare, at around 0.1% per year.

Table 13 Diagnostic accuracy of nodule volume and volume doubling time (NELSON study)

	All nodules			CT within 3 months			CT after 3 months		
VDT <=590 days									
Sensitivity	23/25,	92.00%	(73.9 to 98.9)	15/17,	88.20%	(64.4 to 98.0)	8/8,	100%	(62.8 to 100)
Specificity	360/412,	87.40%	(83.8 to 90.3)	137/178,	77.00%	(70.2 to 82.6)	223/234,	95.30%	(91.7 to 97.4)
PPV	23/75,	30.70%	(21.3 to 41.9)	15/56,	26.80%	(17.5 to 41.0)	8/19,	42.10%	(23.1 to 63.8)
NPV	360/362,	99.40%	(97.9 to 100)	137/139,	98.60%	(94.6 to 99.9)	223/223,	100%	(98.0 to 100)
Volume >=65mm									
Sensitivity	24/25,	96.00%	(78.9 to 100)	16/17,	94.10%	(71.1 to 100)	8/8,	100%	(62.8 to 100)
Specificity	313/412,	76.00%	(71.6 to 79.9)	94/178,	52.80%	(45.5 to 60.0)	219/234,	93.60%	(89.6 to 96.2)
PPV	24/123,	19.50%	(13.4 to 27.5)	16/100,	16.00%	(10.0 to 24.5)	8/23,	34.80%	(18.7 to 55.2)
NPV	313/314,	99.70%	(98.0 to 100)	94/95,	98.90%	(93.7 to 100)	219/219,	100%	(97.9 to 100)
VDT or volume									
Sensitivity	25/25,	100.00%	(84.2 to 100)	17/17,	100.00%	(78.4 to 100)	8/8,	100%	(62.8 to 100)
Specificity	345/412,	83.70%	(79.9 to 87.0)	124/178,	69.70%	(62.5 to 76.0)	221/234,	94.40%	(90.6 to 96.8)
PPV	25/92,	27.20%	(19.1 to 37.1)	17/71,	24.60%	(15.9 to 36.0)	8/21,	38.10%	(20.7 to 59.2)
NPV	345/345,	100.00%	(98.7 to 100)	124/124,	100.00%	(96.4 to 100)	221/221,	100.00%	(97.9 to 100)

3.2.6.6 Summary of CT surveillance evidence

There is currently limited evidence on the clinical accuracy and effectiveness of using CT surveillance to follow-up small or low-risk solid pulmonary nodules. The evidence suggests that appropriate use of volume doubling time and nodule diameter should be associated with high diagnostic accuracy of the entire surveillance schedule, with a sensitivity near 100% (but based on one study [95]). The possibility of slow growing nodules being malignant is not discarded in the BTS guidelines, but determining the numbers of false-negative results (undetected cancers) was not possible for most studies. Existing studies do not distinguish the sensitivity of the different scans within the surveillance schedule. However, the broader evidence reviewed for the BTS guidelines suggests that while more nodules may be identified to have a VDT of <400 days at 3 months than at 12 months, the proportion of those with a malignant diagnosis is likely to be higher at 12 months. [3]

The specificity of CT surveillance is unclear. Some studies demonstrate that a high specificity can be achieved (section 3.2.6), but lower specificity values have been found in other studies, which may be reflective of heterogeneity in the application of CT surveillance in clinical practice.

There is currently no clear evidence on whether CT surveillance leads to tumour progression before detection. In the one study that reported data on this, most cancers detected after surveillance were still at Stage I at time of detection.

3.2.7 Systematic reviews and meta-analyses of pulmonary nodule biopsy methods

Targeted keyword searches of the Endnote database of comparators, other diagnostic tests in the care pathway and surveillance, together with citation searching in Google Scholar, identified five recent (i.e. published after 2010) meta-analyses [98-102] which reported diagnostic accuracy outcomes for nodule biopsy methods (Table 14). Four meta-analyses were excluded for reporting only 'diagnostic yield' outcomes and not sensitivity and specificity. These were Han 2018 et al[103] who reported a meta-analysis comparing radial endobronchial ultrasound and virtual bronchoscopic navigation transbronchial biopsy versus CT-guided transthoracic needle biopsy; Sryma et al 2021,[104] who evaluated Radial Endobronchial Ultrasound (r-EBUS) guided transbronchial cryobiopsy and conventional forceps biopsy; Mondoni 2016 et al who compard Transbronchial needle aspiration with transbronchial biopsy, and Ali et al's 2018[105] meta-analysis on guided bronchoscopy.

Of the five included meta-analyses, all reported on CT-guided percutaneous transthoracic biopsy methods with one meta-analysis[101] additionally reporting results for radial probe endobronchial

ultrasound-guided transbronchial lung biopsy. Given the very limited likelihood of a non-cancerous sample resulting in a cancer diagnosis, all specificity results were, as would be expected, reported as being 1.00 or very close to 1.00. Although two meta-analyses, Liu et al 2020[98] and Yang et al 2014,[100] reported on CT-guided transthoracic needle biopsy methods, Liu et al 2020[98] reported diagnostic accuracy as an outcome but not sensitivity and specificity. Yang et al 2014,[100] searched three bibliographic databases with two independent reviewers screening studies and extracting data. The authors reported a sensitivity of 0.92 (95% CI: 0.88 to 0.95) from pooling six studies (sample size range 28 to 85); four of the six included studies used CT-fluoroscopy and all but one study used core needles. QUADAS-2 was used to evaluate study quality, with all but one of the studies being judged as having low risk of bias and low applicability concerns. However, no details justifying how these judgements were made were presented.

Zhang 2016 et al[102] searched three bibliographic databases; methods for the screening of studies was not reported although two independent reviewers extracted data. The review compared core-needle biopsy with fine needle biopsy, reporting similarly high sensitivities; 15 of the 21 included studies used core needle biopsy (sample size range: 37 to 901) and six were fine needle biopsy studies (sample size range: 32 to 406). QUADAS-2 was used to quality-assess studies, all of which had an unclear risk of bias rating for both index test and reference standard domains. Most studies also had unclear risk of bias ratings for the flow and timing domain.

Yan et al 2017[99] searched five databases for studies of C-Arm Cone-Beam CT-Guided Percutaneous Transthoracic Needle biopsy with two or three independent reviewers screening studies and extracting data. Eight studies were identified (sample size range 35 to 1108) resulting in a pooled sensitivity of 0.96 (95% CI: 0.93 to 0.98). The quality of the included studies was reported as being generally high, as assessed using QUADAS-2, although no details were presented to describe how individual judgements were arrived at.

In the meta-analysis by Zhan et al 2017,[101] four databases were searched and two reviewers independently screened studies and extracted data. The authors compared radial probe endobronchial ultrasound (r-EBUS)-guided transbronchial lung biopsy versus CT-guided percutaneous transthoracic (PTN) biopsy. The old version of QUADAS was used to give studies a quality score out of 14; the maximum score achieved was 8, with most studies having a low score of between 2 and 4. Meta-analyses of 31 studies of r-EBUS-guided transbronchial biopsy and 14 studies of CT-PTN biopsy showed CT-PTN to have the higher sensitivity, but also higher rates of pneumothorax events needing chest tube drainage (1.09% vs 0.48%). Of the other meta-analyses which reported data on pneumothorax events from transthoracic needle biopsies, one reported a rate of 30% overall and 0.02% for events which needed chest tube drainage,[100] and another reported an overall rate of 19%.[98]

Table 14 Meta-analyses of lung nodule biopsy methods published since 2010 which reported diagnostic accuracy outcomes

Review	Biopsy technique(s)	Pooled diagnostic accuracy results	Pooled safety results
Yang 2014[100] 6 studies	CT-guided transthoracic needle	Sensitivity: 0.92 (0.88 to 0.95) I^2=56%, Specificity: 0.98 (0.90 to 1.00) I^2=0%*, LR+: 11.27 (4.2 to 30.6), LR-: 0.1 (0.06 to 0.19)	Pneumothorax rate: 30% (25% to 34%). 7/341 (0.02%) needed chest tube drainage.
Yan 2017[99] 8 studies	C-Arm Cone-Beam CT-Guided Percutaneous Transthoracic Needle	Sensitivity: 0.96 (0.93 to 0.98) I^2=62%, Specificity: 1.00 (0.91 to 1.00) I^2=81%, LR+: 711.2 (9.5 to 53326) I^2=62%, LR-: 0.04 (0.02 to 0.07) I^2=64%.	Pneumothorax rate: only range reported
Liu 2020[98] 25 studies	CT-guided transthoracic needle	Diagnostic accuracy 90% (88% to 93%), I^2=83%. Subgroup results for type of needle, guidance method & lesion size also reported.	Pneumothorax rate: 19% (15% to 24%), I^2=89% Haemoptysis rate: 12% (8% to 15%), I^2=88%
Zhang 2016[102] 21 studies	CT-guided percutaneous core needle (PCN) vs percutaneous fine-needle aspiration (PNA)	PCN: Sensitivity: 0.95 (0.93 to 0.96), I^2=6%. Specificity: 0.99 (0.98 to 1.0), I^2=21%. LR+: 54.7 (28.6 to 104.7). LR-: 0.06 (0.05 to 0.08), I^2=27% PNA: Sensitivity: 0.90 (0.87 to 0.92), I^2=72%, Specificity 0.99 (0.95 to 1.0), LR+: 24.7 (8.9 to 68.9) LR-: 0.14 (0.08 to 0.24)	NE
Zhan 2017,[101] 45 studies	Radial probe endobronchial ultrasound (r-EBUS)-guided transbronchial lung (TBL) vs CT-guided percutaneous transthoracic needle (PTN)	r-EBUS-TBL: Sensitivity 0.69 (0.67 to 0.71), I^2=81% CT-PTN: Sensitivity 0.94 (0.94–0.95), I^2=91%	r-EBUS-TBL: pneumothorax needing chest tube drainage 0.48% (11 out of 2,284). Severe bleeding 0.087% (2 out of 2,284) CT-PTN: Pneumothorax needing chest tube drainage 1.09% (127 out of 11,697). Severe bleeding: 0.32% (36 out of 11,234)

95% confidence intervals in brackets, LR+ Positive likelihood ratio, LR- Negative likelihood ratio, NE Not evaluated, * Text and forest plots results differ

3.3 Further analyses of clinical effectiveness

As noted in Section 3.1.6, no study presented any evidence on the clinical impact of using EarlyCDT Lung. Similarly, no study has reported any evidence on the diagnostic accuracy or clinical impact of using EarlyCDT Lung within the diagnostic pathway, in combination with Brock and Herder risk assessment. To address these issues, we performed a simulation study to examine the potential impact of using EarlyCDT Lung within the diagnostic pathway, and in accordance with BTS guidance.

3.3.1 Methods

Two papers reported complete data on Brock of Herder risk among study participants. The paper by Al-Ameri et al [6] presented a plot of Brock and Herder risk according to nodule status (malignant vs benign) for all participants in the study. The paper by Perandini et al [77] reported a similar figure for

Herder risk only. Data from both figures was digitally extracted to obtain the predicted risks for every participant in these two studies.

To simulate EarlyCDT Lung test results, among people with malignant tumours, a proportion equal to the estimated test sensitivity was randomly assigned to be test-positive. Similarly, among those with benign tumours, a random proportion equal to 1 – specificity was assigned to be test-positive. This made the strong assumption that EarlyCDT Lung test results are independent of Brock and Herder risk (given malignancy status). Sensitivity and specificity estimates were taken from the "high-specificity" EarlyCDT threshold established in Healey et al 2017 [10](Sensitivity 29%, Specificity 98%) and the corresponding "low-specificity" threshold (Sensitivity 49%, Specificity 80%), to simulate test results at both thresholds, adjusted to ensure people positive at "high-specificity" are also positive at "low-specificity".

As the EAG meta-analysis does not support these estimates of diagnostic accuracy we also analysed an alternative "EAG model" with sensitivity and specificity estimates taken from the bivariate meta-analysis (Figure 10) at the same specificity thresholds (Sensitivity 5.1% at 98% specificity; sensitivity 46% at 80% specificity).

Using the simulated EarlyCDT lung data on disease status, and Brock and Herder risk data from the publications, the post-test risk after Brock or Herder assessment and EarlyCDT Lung was calculated using the approach of Healey et al. Briefly, the estimated sensitivity and specificity were combined with the pre-test risk to calculate the post-test PPV, which was taken to be the post-test risk

Using these predicted post-test risks after EarlyCDT Lung assessment, the diagnostic accuracy of Brock alone, Herder alone, Brock with EarlyCDT Lung and Herder with EarlyCDT Lung was calculated at every percentage risk threshold, with results summarised as ROC curves. Brock and Herder risk were analysed separately, as there was no data on the relationship between Brock and Herder risk.

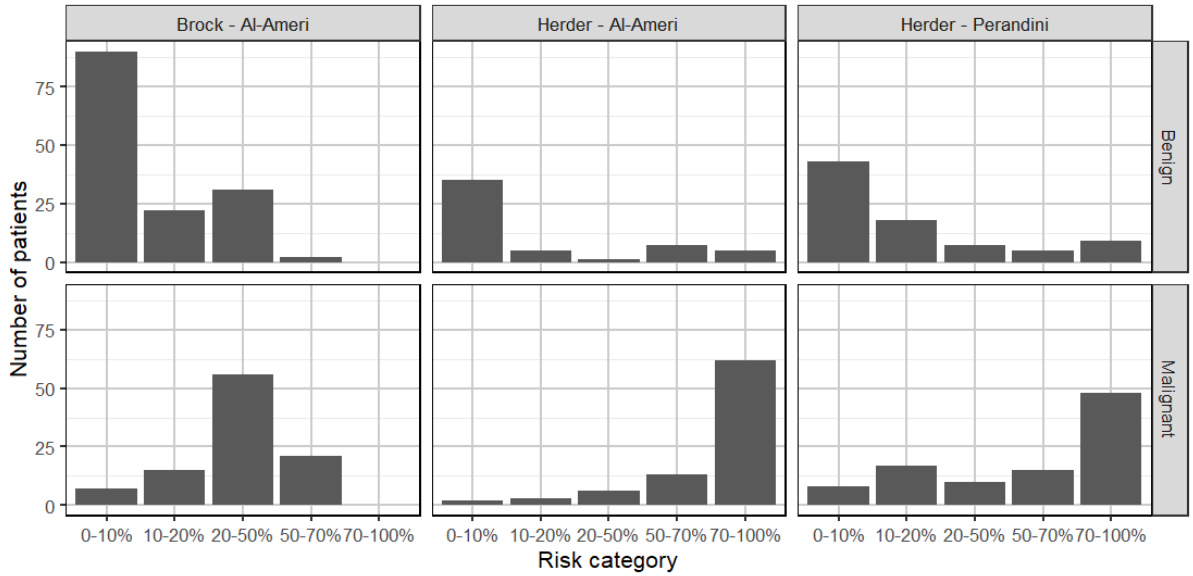
For four arbitrary categories of pre-test risk: 0 - 10% (using Brock risk), and 10 - 20%, 20 - 50% and 50 - 70% (using Herder risk) we used the post-test risk data to calculate the expected percentages of patients who would be correctly and incorrectly reclassified into the next higher risk category or into a risk of over 70%, based on their EarlyCDT Lung results, in order to investigate the clinical impact of adding EarlyCDT Lung to Brock and Herder risk assessment. The 0-10% category corresponds to patients likely to be offered CT surveillance. The other three categories spilt the intermediate risk category (10 -70%) into arbitrary smaller ranges to investigate how EarlyCDT might alter risk (and possibly clinical choices) within the intermediate risk range.

The simulation of EarlyCDT Lung score was repeated 1000 times to obtain a bootstrap sample sufficient to estimate confidence intervals for all results.

3.3.2 Results

Figure 14 presents a bar chart of the extracted data from the Al-Ameri and Perandini publications showing the distribution of risks. People with benign nodules typically have risks below 20% for both Brock and Herder. Results for people with malignant nodules are more variable, with a mix of both low and high risks, although risks with Herder model are skewed towards higher values.

Figure 14 Distribution of risks in Al-Ameri and Perandini studies



3.3.2.1 Diagnostic accuracy of Brock risk

Table 15 summarises the diagnostic accuracy of Brock risk, and combining Brock with EarlyCDT Lung at a risk threshold of 10%, which is the cut-off to distinguish between low-risk nodules to go to CT surveillance and higher risk nodule requiring further investigation. Adding EarlyCDT Lung may slightly improve sensitivity, while reducing specificity, as would be expected because more patients will be "test-positive" after Early CDT Lung assessment. However, changes are small. Similarly changes in in positive or negative predictive values are small. This means that as people with a post-test risk over 10% after EarlyCDT Lung are no more likely to have malignant nodules than when using Brock risk alone. Differences between the Healey and EAG model are small, with no clear evidence of difference. This may be because few patients are test-positive at the "high-specificity" threshold for the Healey model.

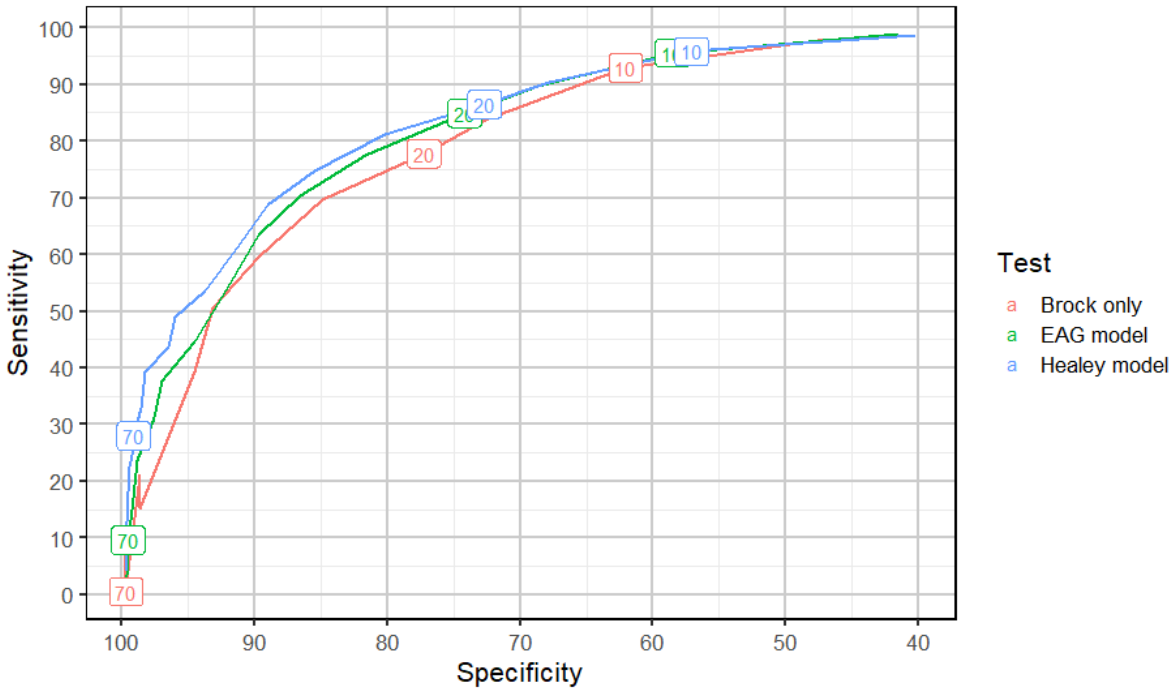
Figure 15 shows the full summary ROC curves at all risk thresholds. Results for thresholds of 10%, 20% and 70% are shown. These results also show no clear benefit of adding EarlyCDT Lung to Brock

risk assessment, with ROC curves being close together. There is a possible, but small, improvement in sensitivity matched by a decline in specificity at each risk threshold. Improvements in sensitivity appear to occur at higher risk thresholds (e.g. 70%), but the Brock risk is not generally used at higher levels of risk. The improvement in sensitivity is notably smaller using the EAG model than the Healey model.

Table 15 Diagnostic accuracy of combining Brock risk with EarlyCDT Lung at the 10% risk threshold

Method	Sensitivity	Specificity	PPV	NPV
Brock model only	92.9%	62.1%	62.6%	92.8%
With EarlyCDT Lung (Healey model)	95.6%	58.5%	61.2%	95.1%
With EarlyCDT Lung (EAG model)	95.8%	57.1%	60.4%	95.2%

Figure 15 Summary ROC curve when combining EarlyCDT Lung with Brock risk



3.3.2.2 Diagnostic accuracy of Herder risk

Table 16 summarises the diagnostic accuracy of Herder risk, and combining Herder with EarlyCDT Lung at risk thresholds of 10% and 70%, which are the cut-offs to distinguish between low-risk, intermediate-risk and high-risk nodules. As for Brock risk, at the 10% threshold adding EarlyCDT Lung to the diagnostic pathway leads to no clear improvement in sensitivity, but with a possible small drop in specificity. Differences in diagnostic accuracy are too small to be conclusive.

At the 70% risk threshold there is a possibility that using EarlyCDT Lung will increase sensitivity substantially while reducing specificity by around 1-2%. This increase in sensitivity is smaller when using the EAG model. This translates into some possible improvement in negative predictive value, but no change in positive predictive value. However, some differences between the Al-Ameri [6] and Perandini [77] data sets makes drawing firm conclusions difficult.

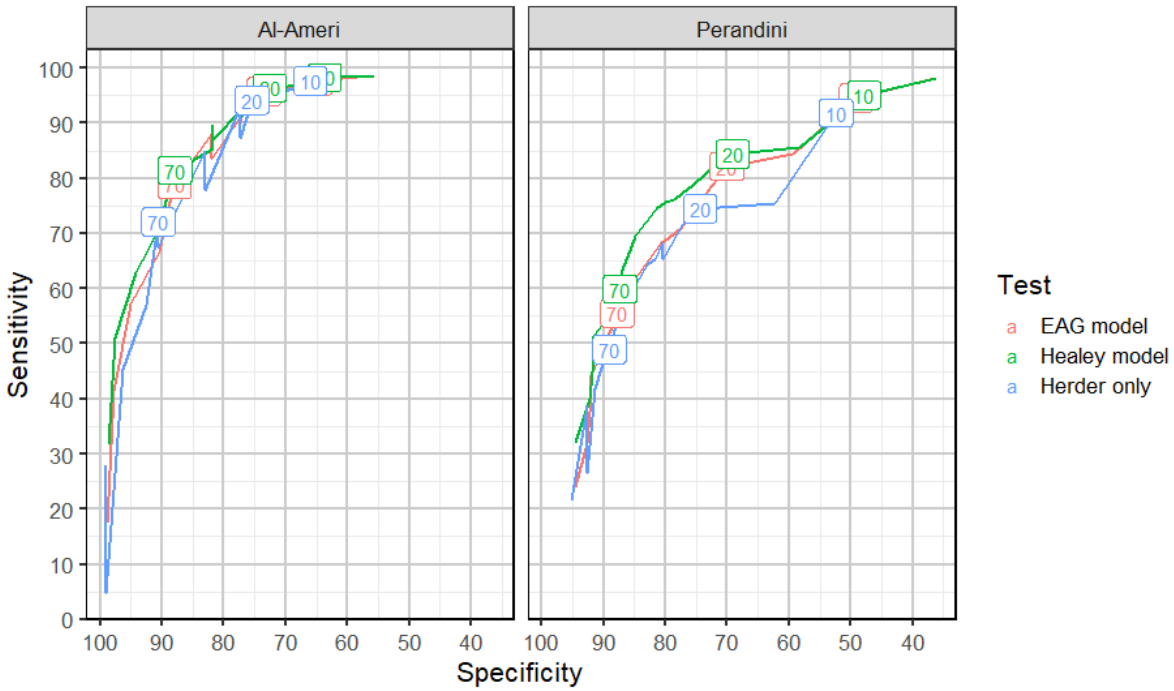
Figure 16 shows the full summary ROC curves at all risk thresholds. Results for thresholds of 10%, 20% and 70% are shown. When using the Healey data and model these results show a possible increase in sensitivity when of adding EarlyCDT Lung to Herder risk assessment, which is most prominent at higher risk thresholds. When using the EAG preferred estimates of diagnostic accuracy, however, this apparent benefit of adding EarlyCDT Lung is substantially reduced.

Table 16 Diagnostic accuracy of combining Herder risk with EarlyCDT Lung at the 10% and 70% risk thresholds

Method	Data	Risk threshold	Sensitivity	Specificity	PPV	NPV
Herder only	Al-Ameri	10	97.7%	66.0%	82.4%	94.6%
With EarlyCDT (Healey model)			98.3%	63.7%	81.5%	95.8%
With EarlyCDT (EAG model)			97.7%	65.3%	82.0%	94.5%
Herder only		70	72.1%	90.6%	92.5%	66.7%
With EarlyCDT (Healey model)			81.4%	87.9%	91.7%	74.6%
With EarlyCDT (EAG model)			78.9%	88.0%	91.4%	72.0%

Herder only	Perandini	10	91.8%	52.4%	69.8%	84.3%
With EarlyCDT (Healey model)			95.1%	47.9%	68.6%	89.2%
With EarlyCDT (EAG model)			94.6%	49.2%	69.0%	88.4%
Herder only		70	49.0%	89.0%	84.2%	59.3%
With EarlyCDT (Healey model)			59.9%	87.3%	85.0%	64.6%
With EarlyCDT (EAG model)			55.6%	87.8%	84.4%	62.3%

Figure 16 Summary ROC curve when combining EarlyCDT Lung with Herder risk



3.3.2.3 Impact on clinical decision making

To assess the impact of adding EarlyCDT Lung to Brock and Herder risk assessment, Table 17 shows how many patients would be reclassified into a higher risk group after using EarlyCDT Lung for four

risk categories (0-10% after Brock risk assessment; 10-20%, 20-50%, 50-70% after Herder risk assessment). Numbers are shown as a percentage of the pre-test risk group. Fuller results, including 95% bootstrap confidence intervals are given in Appendix Table 34.

In the 0-10% risk group, assessed using Brock risk, the numbers of people with a malignant nodule correctly reclassified as having over 10% risk (and so meriting a PET-CT scan or biopsy rather than CT surveillance) is fairly small at around 3% for both Healey and EAG models. This is because of the low diagnostic accuracy of EarlyCDT Lung in this risk group, and the small number of malignant nodules. A small number of people with benign nodules will be incorrectly reclassified as over 10% risk (either 5.1% or 7.5% of the group). The number on incorrect reclassifications exceeds the number of correct reclassifications, by approximately 2:1. The increase in risk with a positive EarlyCDT Lung is never sufficient to increase risk to above 70%.

Therefore, combing Brock with EarlyCDT Lung may mean that more people are wrongly reclassified and will have an unnecessary biopsy than are correctly reclassified. It is unclear whether any benefits of correctly identifying some malignant nodules will outweigh the harms of these unnecessary biopsies, or what the clinical benefits of EarlyCDT Lung might be in smaller nodules where biopsy would not be feasible.

The results in intermediate risk patients, after Herder risk assessment appear more favourable. Using the model proposed by Healey, at 20% pre-test risk or above, 20-35% of patients will be correctly reclassified to a higher risk, generally to over 70% risk. By comparison, few patients with benign nodules are wrongly reclassified (at most 7%). Results with the EAG model were broadly similar. This suggests that a positive EarlyCDT Lung test in this risk range may be a good indicator of a malignant nodule.

There is some variation in results between the Al-Ameri and Perandini data sets, suggesting some uncertainty in the exact proportions of patients who will have risk reclassified after EarlyCDT Lung.

It is currently unclear what the clinical impact of such a reclassification would be, as it is not clear that there is any clinical benefit from proceeding directly to surgery, rather than first receiving a biopsy.

Table 17 Summary of patient risk reclassification when using EarlyCDT Lung in combination with Brock and Herder risk.

Test	Data	Model	Risk group	As proportion of risk group			
				Correctly upgraded	Incorrectly upgraded	Correctly upgraded to >70% risk	Incorrectly upgraded to >70% risk
Brock	Al-Ameri	Healey model	0 to 10%	3.0	7.5	0.0	0.0
		EAG model	0 to 10%	2.8	5.1	0.0	0.0
Herder	Al-Ameri	Healey model	0 to 10%	1.5	3.2	0.0	0.0
			10% to 20%	16.0	12.5	3.4	0.2
			20% to 50%	34.9	0.3	29.0	0.3
			50% to 70%	27.9	6.8	27.9	6.8
		EAG model	0 to 10%	0.0	1.2	0.0	0.0
			10% to 20%	16.4	12.6	0.0	0.0
			20% to 50%	31.8	0.0	1.4	0.0
			50% to 70%	28.6	6.9	28.6	6.9
Herder	Perandini	Healey model	0 to 10%	6.2	7.3	0.0	0.0
			10% to 20%	20.8	10.1	3.2	0.7
			20% to 50%	20.6	4.2	16.4	0.9
			50% to 70%	32.3	4.9	32.3	4.9
		EAG model	0 to 10%	5.2	5.1	0.0	0.0
			10% to 20%	21.1	10.2	0.0	0.0
			20% to 50%	8.6	2.5	0.0	0.0
			50% to 70%	32.8	5.0	32.8	5.0

3.4 Discussion

3.4.1 Key conclusions

The most important conclusion with regard to clinical data on EarlyCDT Lung is that there have been only 5 cohort studies (including 695 patients) of people with pulmonary nodules who have received EarlyCDT Lung. Three of these cohorts are only currently available as conference abstracts. None of the cohorts explicitly performed EarlyCDT Lung after detection of pulmonary nodules using CT scans

of the cohorts, and so none of the cohorts are properly within the BTS guidance pathway for diagnosing pulmonary nodules. Consequently, there are substantial concerns with potential for bias in these cohorts, because the timing of EarlyCDT is not after CT scans. There are also concerns with the lack of independent assessment of EarlyCDT Lung, with only one fully published cohort study that was not funded or conducted by the manufacturer.

Although the evidence is limited, the existing data suggests that EarlyCDT Lung has low diagnostic accuracy to detect cancer in people with pulmonary nodules. Our bivariate analysis suggests a diagnostic accuracy of around 26% sensitivity at 90% specificity. This is notably lower than the sensitivity reported by the manufacturer (e.g. 41.3% sensitivity for 90.6% specificity in Healy et al 2017). Consequently, the predicted increase in risk from having a positive EarlyCDT Lung test is notably lower than the model presented by the manufacturer (see Figure 10). We identified very little evidence on diagnostic accuracy when combining EarlyCDT Lung with other tests, or by nodule size. We identified no published evidence on the clinical impact of using EarlyCDT in patients with pulmonary nodules (such as changes in diagnosis, or in subsequent testing).

We identified few studies of the Brock and Herder risk models for diagnosing pulmonary nodules. The available evidence suggests a high diagnostic accuracy for both tests, with an AUC of 92% (from 8 studies) for Brock model, and an AUC of 84% (from 5 studies) for the Herder model. By comparison, the estimated AUC for EarlyCDT Lung was somewhat lower, at 69.4%. Given the comparatively high diagnostic accuracy for Brock and Herder models compared to Early CDT Lung, it is unclear whether adding EarlyCDT Lung to those tests could substantially improve diagnostic accuracy.

Although several meta-analyses of the use of PET-CT in patients with pulmonary nodules were identified, the studies included in these meta-analyses did not report the performance of PET-CT based on nodule size or on pre-test likelihood of malignancy, as categorised in clinical guidelines. Further searches identified only two studies which stratified results either by pre-test risk or by nodule size.

We identified little evidence on the impact of undergoing CT surveillance. Based on one study, using volume doubling time and nodule diameter had very good diagnostic accuracy to detect malignant nodules. Overall CT surveillance appeared to detect malignant nodules within one year, although there is some uncertainty as to the prevalence and progression of malignant tumours in patients undergoing CT surveillance. It is currently unclear what clinical value using EarlyCDT Lung to remove patients from CT surveillance would offer, partly because the harms of CT surveillance for

small nodules are not well quantified (i.e. the harms that would be avoided by EarlyCDT Lung prompting earlier intervention).

There is adequate evidence providing diagnostic accuracy estimates for CT-guided transthoracic needle biopsy. Better quality studies of radial probe endobronchial ultrasound (r-EBUS)-guided transbronchial lung biopsy may be needed, although they are probably less widely used than CT-guided biopsy.

Simulation studies suggest EarlyCDT Lung is unlikely to offer meaningful clinical improvement for low-risk nodules (0-10%), as adding EarlyCDT Lung to Brock risk appears to result in little change in diagnostic accuracy over using Brock risk alone, with a decline in specificity, but little or no improvement in sensitivity. Using EarlyCDT Lung in patients with low-risk nodules appears to identify few additional genuinely malignant nodules and may lead to more false-positive results than true-positives, and so potentially more people being offered unnecessary biopsies.

EarlyCDT Lung may have some use in identifying malignant nodules among those classified as intermediate risk after Herder risk assessment. At the 70% risk threshold, adding EarlyCDT Lung to Herder may improve the sensitivity for only a small decline in specificity. Consequently, a large proportion of malignant nodules in the intermediate risk group will be correctly identified by EarlyCDT Lung, mostly reclassified to having a new risk of over 70%, with comparatively few false-positive reclassifications. However, these false-positive patients might then needlessly undergo operations with morbidity and mortality risk. It should be noted that these conclusions are from a simulation study, requiring strong modelling assumptions, with high uncertainty. It is also unclear what the clinical benefits to patients would be.

3.4.2 Generalisability

Much of the data on EarlyCDT Lung is either in patients without pulmonary nodules, or comes from studies where EarlyCDT Lung may have been performed before the nodules were identified. No study explicitly using EarlyCDT Lung within the BTS diagnostic pathway currently exists.

Generalisability to the diagnosis of pulmonary nodules identified from CT scans is therefore highly uncertain, as all analysis assumes that diagnostic accuracy in any patient with pulmonary nodules will apply to those identified by CT scans.

Only two cohorts on EarlyCDT have been published in full, from the USA and Germany. These are likely to be generalisable to the UK population, but may have different diagnostic pathways where BTS guidance is not used, which may impact on generalisability of diagnosis using EarlyCDT Lung.

3.4.3 Strengths and limitations

We performed comprehensive search for EarlyCDT Lung studies. This review is therefore likely to have identified all evidence currently published, including all studies reported only as conference abstracts.

To our knowledge this review is the first meta-analysis of all evidence on EarlyCDT Lung, and the first analysis to investigate the possible impact of adding EarlyCDT Lung to Brock and Herder risk assessment.

Overall, analysis was limited by lack of data, with only two fully published studies, and potential for risk of bias and poor generalisability. This meant there was little scope for statistical analysis, and a lack of robustness in results. The EAG considers that the existing evidence is too limited to draw any firm conclusions on the diagnostic accuracy of EarlyCDT Lung.

There is no published evidence on the clinical impact of EarlyCDT Lung. That meant that clinical impact was investigated by a simulation study only, which required strong assumptions of uncertain validity.

3.4.4 Main gaps and limitations in the clinical evidence

The key gap in the evidence is the limited diagnostic accuracy data specifically in patients with pre-diagnosed pulmonary nodules. The EAG concludes that diagnostic accuracy of EarlyCDT is uncertain and potentially at high risk of bias.

Given this, the validity of the risk model proposed by Oncimmune (Figure 2) is uncertain as it is based on potentially biased results from study in patients without pulmonary nodules. The meta-analysis in this review suggests a lower diagnostic accuracy than that used by the company. The EAG considers that a new model properly reflecting diagnostic accuracy in pulmonary model patients is needed. Any new risk model will require independent validation in further cohort studies.

We identified limited evidence on comparator tests in the BTS diagnostic pathway. The diagnostic accuracy of both Brock and Herder models is uncertain, particularly at key risk cut-offs of 10% and 70% risk. Consequently, there is also substantial uncertainty about the diagnostic accuracy when combining these tests with EarlyCDT Lung. The diagnostic accuracy of volume doubling time in CT surveillance is currently limited to one study, so the ability to identify malignant nodules in patient undergoing CT surveillance is uncertain

We identified no published evidence on the clinical impact of using EarlyCDT Lung. Evidence is needed particularly on:

1. Numbers of patients moving from CT surveillance only to PET-CT scan or biopsy after a positive EarlyCDT Lung test, including clinical benefits and harms of this in terms of earlier diagnosis and unnecessary biopsies.
2. Impact on a positive EarlyCDT Lung test in the intermediate (10% to 70%) risk group. Particularly how clinical management might change if risk is increased but remains within this intermediate range.
3. Impact of moving risk from intermediate to high risk (over 70%) after a positive EarlyCDT Lung test. Whether this would this lead to immediate excision without biopsy and the clinical benefits and risk of excision without biopsy.

There is generally limited evidence on the implementation of the overall BTS pathway, including on patient outcomes. Evidence is needed on:

1. Prevalence of malignancies by tumour size, Brock and Herder risk and their correlations
2. The clinical outcomes for patients undergoing CT surveillance, including time to identify malignant nodules, and disease progression during CT surveillance.
3. Evidence on clinical management choices for patients at intermediate risk, including impact of choosing between CT surveillance, image-guided biopsy and immediate excision or surgery.

4 EVIDENCE ON THE COST EFFECTIVENESS OF EARLYCDT LUNG

This section provides an overview of existing cost-effectiveness evidence on the use of EarlyCDT Lung for the assessment of solid pulmonary nodules, so as to ascertain its generalisability to the relevant decision problem. The review also aimed to identify i) key structural and parameter assumptions, and ii) components of value of the technology, as well as characterise evidence linkage mechanisms used to link these components of value to final outcomes, in the existing cost-effectiveness models.

4.1 Search and studies identified

The search detailed in Section 2.2.1 identified 3,233 record. The first stage of screening identified two potentially relevant records, based on their title and/or abstract. The corresponding full text articles were retrieved and assess for inclusion. The two studies[106, 107] met the inclusion criteria (see Section 2.3.1.2), and were included in this review.

4.2 Methods and key assumptions of the identified studies

The two identified studies are summarised in Table 18. The quality assessment of these studies followed a checklist specific to model-based economic evaluations of diagnostic tests[16] which is reported in Appendix 11.7 (Table 35 and Table 36 for Edelsberg et al. 2018[106] and Sutton et al. 2020,[107] respectively).

Table 18 Summary of cost-effectiveness studies of EarlyCDT Lung

Study, perspective	Population	Population characteristics	Diagnostic comparators	Analytical approach, time horizon	Outcomes
Edelsberg, 2018, US Healthcare system	Patients with incidentally detected intermediate-risk nodules of 8–30 mm and intermediate risk (5–60%) risk of lung cancer	Mean age: 65.3 years % Female: 47.1% % smokers:76.5% %NSCLC/SCLC: 94%/4% Malignancy prevalence: 9.5% Baseline cancer stage distribution for malignant nodules: 100% local	• EarlyCDT Lung • CT surveillance alone . EarlyCDT Lung is a one-off test, while CT surveillance is a repeated at 4, 10, and 21 months. . Patients with a positive EarlyCDT Lung result receive a diagnostic biopsy or wedge resection. Patients with negative tests results either enter or remain in CT surveillance until they test positive or the surveillance interval elapses. It is unclear how patients who test positive to CT surveillance are managed. . Two scenarios evaluate alternative diagnostic accuracy values for EarlyCDT Lung (scenario A: sensitivity 0.41 and specificity 0.93; scenario B: sensitivity 0.28 specificity 0.98)	Decision analytical model, life-time (given use of life-expectancies) Structure not described	Cost per life-year gained and cost per QALY gained Disease stage distribution % stage shift
Sutton, 2020, UK Healthcare provider	Patients with IPNs identified by imaging, which are between 4mm and 20mm in size and carry a risk of malignancy of 10–65% (lung cancer)	Mean age: 62 years Malignancy prevalence: 9.5% Baseline cancer stage distribution for malignant nodules: 87.5% local; 12.5% regional	• EarlyCDT Lung • CT surveillance alone . EarlyCDT Lung is a one-off test, while CT surveillance is a repeated at 3, 12, and 24 months. . Patients with positive tests in either strategy are subject to diagnostic biopsy, followed by surgical removal if the nodule is confirmed to be malignant (or if benign but with nodule growth). . Patients with negative tests results either enter or remain in CT surveillance until they test positive or the surveillance interval elapses. . Two scenarios evaluate alternative diagnostic accuracy values for EarlyCDT Lung (scenario A: sensitivity 0.41 and specificity 0.93; scenario B: sensitivity 0.28 specificity 0.98)	Decision analytic model: - Decision tree + - Markov model; multiple health states (undiagnosed benign, diagnosed benign, undiagnosed local, undiagnosed regional, undiagnosed distant, diagnosed local, diagnosed regional, diagnosed distant, recurrence mortality, disease free, cancer mortality) Life-time horizon	Cost per QALY EVPI EVPPI

EVPPI, expected value of partially perfect information; EVPI, expected value of perfect information; IPNs, indeterminate pulmonary nodules; NSCLC, non-small cell lung cancer; SCLC, small cell lung cancer

Both studies assess the cost-effectiveness of EarlyCDT Lung compared to routine CT surveillance for the diagnosis of lung cancer in patients with solid pulmonary nodules using a decision modelling approach. The two studies took a healthcare payer perspective; Edelsberg et al., 2018,[106]is set in the US healthcare system, while Sutton et al., 2020 is set in the UK National Health System (NHS). Edelsberg et al., 2018,[106] assumed a cost higher cost of EarlyCDT Lung than Sutton et al., 2020[107] (cost per test: $575 vs. £70).

The proposed positioning of EarlyCDT Lung and target patient population is defined differently in the two studies. The target population in Edelsberg et al., 2018, [106]is defined as patients with incidentally detected nodules of 8–30 mm and intermediate risk (5–60%) risk of lung cancer. This population was considered relevant by the authors, because there was some evidence that this group of patients might be followed up with routine CT surveillance instead of PET-CT, as recommended by American College of Chest Physicians (ACCP) for nodules of this size. Sutton et al., 2020,[107] considered a patient population with nodules with a size 4-20mm, and a risk of malignancy (lung cancer) of 10–65%. The population choice was not explicitly justified in this study. The authors state only that the BTS guidelines consider that some nodules with a 10-70% risk of malignancy have too low a risk to be considered for biopsy and can instead be followed up with CT surveillance or 'watchful waiting'.

Both studies assumed the same prevalence of malignant nodules (9.5%) sourced from a study on the diagnostic follow-up and management of nodules of 8-30mm size by a pulmonologist and/or a thoracic surgeon in a North American (US and Canada) setting.[108] In what concerns the cancer stage distribution at baseline, Edelsberg et al., 2018, [106] assumed all patients with malignant nodules had local disease (sourced from Tanner et al., 2017[108]), while Sutton et al., 2020,[107] assumed only 87.5% with local disease and the rest regional disease (sourced from Gould et al., 2003[109]).

Both studies compare EarlyCDT Lung in addition to a CT surveillance schedule versus CT surveillance alone. Patients in the EarlyCDT Lung + CT surveillance strategy receive a one-off EarlyCDT lung test at the start of the model, which produces a dichotomous test result (positive/negative). Neither of the studies report the criteria for positivity (i.e., the diagnostic cut-off for each of the seven autoantibodies in the test panel and whether one or more autoantibodies levels need to be elevated for the test to be positive). Patients who test positive with EarlyCDT Lung receive a diagnostic biopsy (some patients in Edelsberg et al., 2018,[106] receive wedge resection instead). Biopsy is assumed 100% accurate and patients with benign nodules (false positive to EarlyCDT Lung) proceed to CT surveillance while those with malignant nodules proceed to excision. Patients who test negative with EarlyCDT Lung enter CT surveillance to assess tumour growth, following the schedules described in Table 18.

All patients under CT surveillance remain until they test positive (i.e., until nodule volume doubles in Edelsberg et al., 2018;[106] unclear in Sutton et al., 2020[107] but also based in change in nodule volume) or reach the end of the surveillance interval.

Patients who test positive during CT surveillance in Sutton et al., 2020,[107] receive a biopsy to excise the nodule (type of surgery not specified in the manuscript). It is unclear how patients who test positive during CT surveillance in Edelsberg et al., 2018,[106] are managed.

The modelling approach taken by Edelsberg et al., 2018,[106] is insufficiently described in the manuscript, but it appears to quantify long-term outcomes using life-expectancy projections (and HRQoL gains) conditional on nodule malignancy, cancer histology (non-small cell lung cancer [NSCLC] or small cell lung cancer [SCLC]), cancer stage (local, regional or distant), and patient characteristics (see Section 4.4.3). The model tracks volume doubling time (VDT) and cancer stage progression over 2 years, with VDT and progression probabilities informed by Gould et al., 2003.[109] The stage distribution for each strategy is calculated at 2 years based on this. It is unclear from which time point in the model were the life expectancy projections applied. This model also considers overdiagnosis of indolent malignant nodules (i.e., nodules which are not aggressive despite being malignant); this was implemented as a reduction in the malignancy prevalence of 18% (based on data from a lung cancer screening population). The authors state that since all lung cancers are diagnosed by both strategies under comparison, overdiagnosis is equally frequent for these strategies and it only affects the *de facto* prevalence.

The modelling structure in Sutton et al., 2020,[107] comprises a decision tree and a Markov model with monthly cycles and half-cycle correction. The Markov model component is stated to have the same structure as the model used by Gould et al., 2003. [109] At the first cycle in the model, the decision tree dichotomises patients in each strategy according to the disease status (positive/negative) and then applies test diagnostic accuracy estimates to classify patients according to their test results as true positive (TP), false negative (FN), true negative (TN) and false positive (FP). Patients then enter the Markov component of the model according to whether they have been correctly diagnosed (diagnosed health states: benign or malign) or not (undiagnosed health states), and their disease stage for patients who have malignant nodules.

In the Markov model, only patients in undiagnosed cancer health states (local or regional) seem to be able to progress between disease stages (local to regional and regional to distant). Progression between cancer stages for undiagnosed is dependent on nodule growth over time. Patients in the diagnosed benign state remain in the state but may undergo surgical biopsy in future cycles if CT surveillance detects nodule growth. Patients with diagnosed (and treated with either surgery alone or

with chemotherapy or radiotherapy) malignant local and regional nodules have a time-dependent mortality risk due to cancer (recurrent or regional for local and regional cancer, respectively) for five years, after which they transition to the disease-free state. Patients in the diagnosed distant states have a lifetime constant risk of cancer related mortality, as do patients with undiagnosed malignant nodules. Patient in all health states are subject to age adjusted general population mortality. Cancer related mortality, probability of benign nodule growth and disease progression probability were sourced from Gould et al., 2003. [109]

Two scenarios are analysed separately in each study considering alternative values for the diagnostic accuracy of EarlyCDT Lung: scenario A considers a sensitivity and specificity of EarlyCDT Lung of 41% and 93%, respectively, and scenario B a sensitivity and specificity of 98% and 28%, respectively (both value sets are sourced from Healey et al., 2017,[10] see details Section 3.1). These scenarios were each evaluated as a pair-wise comparison against CT surveillance alone. One of the key differences between the two studies is that CT surveillance is assumed to detect all malignant nodules over the two years follow-up in Edelsberg et al., 2018, while in Sutton et al., 2020[107] there is misclassification under CT surveillance leading to a proportion of undiagnosed malignant nodules at the end of the surveillance schedule. Sutton et al., 2020,[107] synthesised diagnostic accuracy data from studies identified in Gould et al., 2003[109] to inform the sensitivity and specificity of CT scan (92.3% and 72.3%, respectively). Both Edelsberg et al., 2018,[106] and Sutton et al., 2020, [107] assumed that biopsy was 100% accurate.

4.3 Results of the identified studies

Table 19 summarises the cost-effectiveness results from the two studies. Both studies conclude that EarlyCDT Lung is a cost-effective use of health payer resources compared to CT surveillance alone, as the incremental cost-effectiveness ratios (ICERs) are below the cost-effectiveness thresholds in the studies' jurisdiction. Despite the two models relying on similar data sources and assumptions, the ICERs of EarlyCDT Lung + CT surveillance vs. CT surveillance differ substantially across the two studies.

Table 19 Summary of cost-effectiveness results in the studies of EarlyCDT Lung

	Total cost	Incremental cost*	Total QALYs	Incremental QALYs*	Total LYG	Incremental LYG*	ICER* (per QALY)
Edelsberg et al., 2018							
CT surveillance	$4,040	-	9.793	-	12.130	-	-
EarlyCDT Lung A+CT surveillance	$4,989	$949	9.832	0.039	12.183	0.053	$24,330
EarlyCDT Lung B+CT surveillance	$4,722	$682	9.821	0.027	12.167	0.037	$24,831
Sutton et al., 2020							
CT surveillance	£2,261	-	10.685	-	-	-	-
EarlyCDT Lung A+CT surveillance	£2,410	£149	10.7465	0.0614	-	-	£2,417
EarlyCDT Lung B+CT surveillance	£2,358	£97	10.7308	0.0457	-	-	£2,121

*compared to CT surveillance; A, sensitivity/specificity = 0.41/0.93; B, sensitivity/specificity = 0.28/0.98; LYG, life years gained

Beyond the difference in the per patient costs of the test itself, it is difficult to understand which parameters are driving these differences in cost-effectiveness given the lack of detail and clarity on important analytical choices in both models. Furthermore, Sutton et al., 2020, [107] does not report life-years gained in the model or conduct a thorough exploration of parameter and structural uncertainty, which would have aided interpretation of differences between models. The main differences between the two models in terms of parameterisation and structural assumptions are:

- Higher cost of EarlyCDT Lung in Edelsberg et al., 2018,[106] compared to Sutton et al., 2020[107](cost per test: $575 vs. £70);
- Higher costs of health care in Edelsberg et al., 2018;[106]
- Baseline distribution of malignant nodules across disease stages;
- Actual malignancy prevalence estimate applied in the model – approximately 2% lower in Edelsberg et al., 2018;[106]
- Assumption of no misclassification at the end of surveillance in Edelsberg et al., 2018;[106]
- Edelsberg et al., 2018,[106] only explicitly models disease progression during the 2 years of CT surveillance with a 55.3% probability of progression over 2 years (or an annual probability of 33.1%);

- Modelling of long-term outcomes as life-expectancy (HRQoL adjusted and unadjusted) pay-offs[106] vs. through a Markov model.[107]

We do not focus here on examining the studies' results, as these are unlikely to be appropriate to inform the decision problem defined by the NICE scope (see Section 4.4.1). However, the differences in results between studies, suggest that there are differences in terms of how each study modelled the value of EarlyCDT Lung and the evidence linkage used to translate this into impact on health and costs differences. In the next section, we first critique the two studies in terms of their relevance to the scope of this assessment. We then examine in more detail the drivers and components of value of EarlyCDT Lung (vs. CT surveillance) and the evidence linkage approach taken, to better understand the modelling of the mechanisms of value accrual and support the development of a conceptual model to assess the cost-effectiveness EarlyCDT Lung.

4.4 Critique

4.4.1 Decision problem and relevance to NICE DAR scope

The suitability of the identified studies to inform the decision problem defined by the NICE DAR scope is assessed in this section. Table 20 compares how the studies relate to the NICE scope in three key areas where the EAG identified a lack of alignment.

Table 20 CE studies of EarlyCDT Lung vs. scope

	NICE DAR scope		**Edelsberg et al., 2018**	**Sutton et al., 2020**
Patient population	Patients without previous history of cancer and with solid pulmonary nodules (SPN) >5mm in diameter or >80mm^3 in volume		Patients with incidentally detected intermediate-risk nodules of 8–30 mm and intermediate risk (5–60%) risk of lung cancer	Patients with nodules size 4-20mm, and a risk of lung cancer of 10–65%.
Position in the pathway	Multiple positions: 1. Nodules 5-8mm in diameter or 80-300mm^3 in volume 2. Nodules >8mm in diameter or >300mm^3 in volume with ○ 2.1. <10% risk of malignancy using the Brock model ○ 2.2. ≥10% risk of malignancy using the Brock model 3. Nodules >8mm in diameter or >300mm^3 in volume with ○ 3.1.<10% risk of malignancy using the Herder model ○ 3.2. 10%- 70% risk of malignancy using the Herder model	Current practice: 1. CT surveillance 2.1. CT surveillance 2.2. PET-CT 3.1. CT surveillance 3.2. Image guided	Patients assumed to be eligible for PET-CT, but who do not receive this test Comparator: CT surveillance	Unclear, data sources suggest similar to Edelsberg et al, 2018 Comparator: CT surveillance

		biopsy, excision biopsy or CT surveillance		
Test result format and use of test result	Categorical test result: low, moderate, high Upgrade patient pre-test malignancy risk scores		Binary test result: positive (malignant), negative (benign) Identify malignancy	

The study populations in the identified studies do not appear to match the population defined in the scope to the current assessment. Malignancy prevalence, a key model parameter, is informed in both studies by data the Tanner study,[108] a US study which includes patients with pulmonary nodules at an intermediate risk of malignancy who were managed with CT surveillance despite indication for PET-CT scan according to ACCP guidelines. The ACCP guidelines for the management of pulmonary nodules are not followed in UK clinical practice, and differ from corresponding BTS guidelines in how to perform malignancy risk assessment and the risk cut-offs used to guide diagnostic follow-up (see Section 1.1.1). Furthermore, clinical opinion suggests that adherence to BTS guidelines is high and that PET-CT is widely available in UK clinical practice, so in the UK patients with nodules at an intermediate risk of malignancy would receive PET-CT and further risk assessment, rather than going directly to CT surveillance at the first stage of risk assessment with the Brock model. Since the Edelsberg et al., 2018 study[106] is set in the US healthcare system and they are explicitly trying to evaluate the use of EarlyCDT Lung where clinical guidance is not adhered to, the data from Tanner et al., 2017[108] may be of some relevance. In Sutton et al., 2020,[107]which is set in the UK NHS, the authors do not justify the selection of this study to inform malignancy prevalence.

It is unlikely that the characteristics of patients in Tanner et al., 2017,[108] are comparable to those of the patients in the current assessment population. The EAG considers that the prevalence estimates sourced from this study, and used by both cost-effectiveness studies, are unlikely to be of relevance to the populations defined in the NICE DAR scope.

Both studies compare EarlyCDT Lung in addition to CT surveillance. These studies do not discuss other diagnostic comparators, or alternative positioning of the new technology in the diagnostic pathway. In Sutton et al., 2020, it is not even clear where exactly in the diagnostic pathway is the technology being used, given that there are two points for risk assessment (pre-PET-CT with the Brock model and post-PET-CT with the Herder model), and that the nodules in this study are in a category of risk (10-65%) which does not match those defined by the BTS guidelines. The closest match for the patients in Sutton et al., 2020, would appear to be to patients with intermediate risk nodules (10-70%) following assessment with PET-CT and the Herder model (even if the evidence

used to populate the model is not necessarily reflective of this group). In this position in the diagnostic pathway follow-up options include CT surveillance, but also imaging guided biopsy or excision biopsy. This suggests that not all relevant comparators have been considered in this study.

The diagnostic accuracy of EarlyCDT Lung is not modelled as proposed in the information submitted by the company in either of the identified studies. In Edelsberg et al., 2018,[106] and Sutton et al., 2020,[107] EarlyCDT Lung diagnostic accuracy reflects its ability to correctly identify malignancy, whereas the company proposes that EarlyCDT Lung results are used to update patient pre-test malignancy risk scores according to a risk calculator and inform clinical decision based on the updated score (see Section 3.1.7.3). The EAG also notes that the diagnostic accuracy of EarlyCDT Lung is likely to be overestimated in Healey et al., 2017 [10] (see Section 3.1.7.3), which could bias the cost-effectiveness results of the identified studies.

The EAG concludes that the existing studies cannot directly inform the current decision problem, given the substantial differences between the models and the NICE DAR scope. The EAG concerns in regards the suitability of the studies to inform the decision problem, stem from the following issues:

i. The studies population is unlikely to be reflective of patients in the UK clinical practice at any point in the diagnostic pathway, and estimates of prevalence lack generalisability to the population of interest.
ii. The position of EarlyCDT Lung in the diagnostic pathway as modelled in these studies does not match the potential uses of the technology under the defined scope and the diagnostic comparators considered do not include all relevant alternatives.
iii. The diagnostic test use in the studies does not match the use proposed by the company in the DAR. The diagnostic accuracy metric (specificity and sensitivity at a single diagnostic threshold) of the evidence used in the studies is not appropriate to inform the diagnostic accuracy of EarlyCDT Lung used as part of malignancy risk assessment.

4.4.2 Components of value

In this section we examine the components of value (i.e., the features of the test in regards to comparators that allow establishing and quantifying trade-offs, the balance of which determines the net value of the technology) modelled in each study and how the evidence on these was linked to health and cost outcomes. The components of value for EarlyCDT Lung in relation to CT surveillance identified across the two studies are summarised in Table 21.

Table 21 Components of value for EarlyCDT Lung in relation to CT surveillance

	Components of value for EarlyCDT Lung in relation to current practice (routine CT surveillance) considered in Sutton 2020 and Edelsberg 2018
1	Additional cost of EarlyCDT Lung test in all individuals
2	Improved outcomes from early/increased detection of lung cancer (stage shift) in true positives to EarlyCDT Lung
3	Additional costs and risk of adverse events of further investigations on positives to EarlyCDT Lung
4	Avoided costs of CT surveillance in all positives to EarlyCDT Lung

Both models consider the additional cost of EarlyCDT Lung compared to CT surveillance alone (item 1, Table 21), although the cost per test is substantially higher in Edelsberg et al. 2018,[106] compared to Sutton et al., 2020[107] ($575 vs. £70).) The two test cost estimates were informed by Oncimmune, and neither study details how these estimates were calculated or whether they include only the cost of the test or also other associated costs (e.g., training and administration costs). It is unclear why there is such a difference in this parameter between the two studies. It is worth noticing that the EarlyCDT cost per test included in the within-trial cost-effectiveness analysis of EarlyCDT Lung in the context of screening in Scotland (ECLS [see Section 3.1.7.1]) the cost per test as informed by Oncimmune was £95 (per blood test, based on $124 per kit).[13] This study further included a cost for blood collection, consisting the cost of 15 minutes nurse time at the GP practice. As noted in Section 1.3.1.1, the cost of EarlyCDT Lung testing should include not only the cost of the test, but also those of i) consumables required to process the test, ii) test administration (including blood collection), iii) training needed to process/administer the test, and iv) costs of delivering test results to individuals. Both Edelsberg et al. 2018,[106] and Sutton et al., 2020[107] may not have included all relevant categories of cost in the cost of EarlyCDT Lung testing.

Remaining effects are indirect, in that the impact of the test on outcomes is realised indirectly by tailoring patient management to the test result in each individual. The studies present a common and key value mechanism for EarlyCDT Lung compared to CT surveillance: they establish a link between early diagnosis of lung cancer and improved health outcomes for patients who have a true positive result to EarlyCDT Lung (item 2, Table 21). The mechanism by which this improvement is achieved is via a cancer "stage shift", whereby patients diagnosed earlier are assumed to be in earlier stages of the disease and therefore have a better prognosis from treatment. The mechanism of value from increased detection is also expressed as early detection and assumes that cancers missed by CT surveillance would present clinically later in time. Increased detection with EarlyCDT Lung is only modelled in Sutton et al., 2020, [107] where having one additional test in the strategy leads to an increased yield of true positive results for the overall strategy of EarlyCDT Lung (followed by CT

surveillance for the negatives or biopsy for the positives) compared to CT surveillance alone, as CT surveillance is not assumed to be a perfect test. In Edelsberg et al., 2018,[106] this value component is not captured, because when CT surveillance is assumed 100% accurate, there is no defence between strategies in the number of correctly identified malignant tumours.

Both studies also include a cost and mortality impact from biopsies for positive results with EarlyCDT Lung (item 3, Table 21), as all EarlyCDT Lung positive results are assumed to require a follow-up with biopsy. Sutton et al., 2020,[107] further considers the disutility associated with biopsies, although it is unclear how this was applied in the model.

Although Edelsberg et al., 2018, [106] explicitly models the impact of overdiagnosis of indolent malignant tumours, this is not reflected as a value driver for EarlyCDT Lung as it equally impacted both strategies under comparison. The authors justify this approach based on their assumption that all lung cancers are correctly identified in the model, and therefore overdiagnosis would be the same for both strategies. However, the authors do not comment that CT surveillance should be able to differentiate between indolent and aggressive nodules, as the former would not grow at the same rate as aggressive nodules. Indolent nodules should be less likely to be overdiagnosed under CT surveillance. Thus, both Edelsberg et al., 2018,[106] and Sutton et al., 2020[107] (which makes no attempt to model this) miss a potential value component for EarlyCDT Lung.

4.4.3 Evidence linkage

Table 22 illustrates how the value components of EarlyCDT Lung were modelled with a focus on the evidence linkage approach taken to connect the patient classification based on test results to clinical decisions and these to patient final outcomes, in accordance to the framework proposed by Soares et al., 2018.[23] The table details, for each testing strategy, the alternative diagnostic pathways that patients can follow based on the sequence of tests and their results, whether patients can be misclassified by the overall test sequence and the final classification of patients at the end of the sequence. It then lists the treatment choice for the different classification. The table also summarises the mechanism of linking patient classification to model outcomes, by making explicit the conditional relationships in the model.

In both models, EarlyCDT Lung is administered once at the start of the diagnostic pathway. CT surveillance consists of repeat CT scans which measure tumour growth between scans for all test sequences. For simplicity, in Table 22 CT surveillance is represented as a single test (CTsurv) in the test sequences and its test result is indicated as negative (-) if none of the CT scans in the sequence has a positive result and as positive (+) if one of the CT scans in the sequence has a positive result (ending the surveillance).

Table 22 Evidence linkage mechanism between classification, treatment choices and outcomes

Study	Pathways of test sequences	Misclass*	Final classification (diagnosis)	Treatment choice	longer term outcomes | diagnostic workup & treatment
Edelsberg 2018	EarlyCDT Lung(-) → CTsurv(-)	No	benign or malignant	Treatment for malignant No treatment for benign	**Direct effects** costs| tests; mortality| biopsy **Indirect effects** cancer stage| time to diagnosis mortality| malignancy/treatment, cancer stage, smoking status, age HRQoL| malignancy, cancer stage, cancer histology, age** costs| malignancy/treatment
	EarlyCDT Lung(-) → CTsurv(+) → Biopsy(+)	No	malignant		
	EarlyCDT Lung(+) → Biopsy(+)	No	benign		
	EarlyCDT Lung(+) → Biopsy(-)	No	benign		
	CTsurv(-)	No	benign or malignant		
	CTsurv(+)→Biopsy(+)	No	malignant		
	CTsurv(+) →Biopsy(-)	No	benign		
Sutton 2020	EarlyCDT Lung(-) → CTsurv(-)	Yes	undiagnosed (benign or malignant)	No treatment	**Direct effect:** costs| biopsy, probability of biopsy complications, positive biopsy result; mortality| probability of biopsy; HRQoL| probability of biopsy complications **Indirect effects** cancer stage| probability of progression while undiagnosed mortality| malignancy, detection/treatment, cancer stage, time on health state, age HRQoL| malignancy, cancer stage, age costs| treatment, probability of surgical complications
	EarlyCDT Lung(-) → CTsurv(+) → Biopsy(+)	No	malignant	Surgery	
	EarlyCDT Lung(-) → CTsurv(+) → Biopsy(-)	No	benign with growth	Surgical biopsy	
	EarlyCDT Lung(+) → Biopsy(+)	No	malignant	Surgical biopsy	
	EarlyCDT Lung(+) → Biopsy (-) → CTsurv(-)	No	benign with/without growth	No treatment/surgical biopsy	
	CTsurv(-)	Yes	undiagnosed (benign or malignant)	No treatment	
	CTsurv(+) →Biopsy(+)	No	malignant	Surgery	
	CTsurv(+) → Biopsy(-)→ CTsurv(-)	No	benign with/without growth	Surgical biopsy	

*This column captures whether misclassification is possible (Yes/No) in each diagnostic pathway defined by the test sequences; **, for patients with benign tumours only.
CTsurv, CT surveillance; HRQoL, health-related quality of life; VDT, volume doubling time.

4.4.3.1 Evidence linkage in Edelsberg et al., 2018

Patients in this model are all assumed to be correctly diagnosed as having a benign or malignant nodule at the end of the test sequence for all strategies, as the last test (CT surveillance or biopsy) in every sequence is assumed to be a perfect test. Since there are no misclassified patients in the model all nodules are appropriately treated; benign nodules receive no treatment and malignant nodules receive cancer treatment (exact treatment not specified).

EarlyCDT Lung impacts on outcomes via the increased use of biopsy to confirm positive results subsequent to EarlyCDT Lung; this includes the added costs of the biopsy and its associated mortality risk.

Long term effects of EarlyCDT Lung are promoted by earlier diagnosis and associated stage shift at diagnosis (section 4.4.2). All malignant nodules are assumed to be at the earliest disease stage (local disease) when they enter the model. The use of EarlyCDT Lung will result in a higher number of malignant nodules being detected at the local stage (out of local, regional and distant). The extent of stage shift is conditional on nodule growth (VDT), although it is unclear how the VDT data was used to estimate probability of progression given volume doubling over time. The authors appear to have used the same observed data on nodule growth from a 1973 study on 67 nodules is detected with chest radiography [110] as per a previous cost-effectiveness study [109], but the assumptions relating volume growth and disease progression are not reported. The EAG notes that Steele et al.,1973[110] predates CT imaging and used a different imaging technique, chest radiography or X-ray, is used to determine nodule size. Chest radiography has worse spatial resolution and a higher threshold for detection of nodules than CT imaging. [109] It highly uncertain whether tumour growth rates derived from chest radiography measurements is suitable to inform growth rates during CT surveillance, especially for smaller nodules (<2cm[109]). The sample size of this study is small (n=67), which also contributes to the uncertainty surrounding this evidence.

The model estimates life-expectancy for patients with malignant tumours conditional on the disease stage, and age. The authors only state that they combined *"data from The Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute and data on relative survival from the National Cancer Database"* to inform life-expectancy, but it is unclear how exactly was lung cancer these data were applied in the model, and whether this is reflective of all malignant nodules receiving cancer treatment in the model. In addition to cancer mortality, the model also considered the other cause mortality of lung cancer patients adjusted for their smoking status. The life expectancy of patients with benign nodules is stated to be based on data from the NLST,[111] but no details are provided on how these estimates were derived. The study does not report at which point in the model the projected life-expectancy estimates are applied.

For patients with malignant tumours, life-time quality-adjusted life years (QALYs) estimates were calculated by applying health state utility values reflecting the cancer stage and its histological type (NSCLC or SCLC) and sourced from the NLST. The HRQoL of patients with benign tumours is assumed to be age-specific and is sourced from published literature.

The model considers the cost of cancer treatment for patients with malignant tumours. This cost is not dependent on disease stage and, therefore, does not rely on a link to disease progression. It is unclear whether this cost includes any long-term costs of the disease or the immediate costs of treating cancer after diagnosis.

4.4.3.2 Evidence linkage in Sutton et al., 2020

The diagnostic pathways in the model by Sutton et al., 2020, [107] are structured differently from those in Edelsberg et al., 2018.[106] This is due to misclassification of nodules being possible with CT scan and to benign nodules with growth detected during CT surveillance being managed with surgical biopsy. in Sutton et al., 2020,[107] patients who test negative on a biopsy following a positive result to EarlyCDT Lung or CT scan are placed under CT surveillance, whereas Edelsberg et al., 2018,[106] does not explicitly state how these patients are managed. The final classification of nodules in Sutton et al., 2020,[107] can be: diagnosed benign with or without growth (TN); undiagnosed benign or malignant (TN and FN), diagnosed malignant (TP).

Sutton et al., 2020[107] considers impacts on cost, mortality and HRQoL via the increased use of biopsy to confirm positive results subsequent to EarlyCDT Lung. Procedural complications of the biopsy are considered both in terms of both cost and disutility. Patients with a positive biopsy result also incur the cost of one appointment with a multidisciplinary team.

Similarly, to the previous model, in Sutton et al., 2020, [107] the longer-term impact on outcomes of EarlyCDT Lung are mediated via an effect on disease progression (reflected on the stage of the disease at diagnose), but the modelling approach taken is different. This model explicitly uses a Markov model to track disease progression for undiagnosed malignant nodules and diagnosed malignant nodules at a regional stage, while progression is assumed to be halted for diagnosed malignant nodules at a local stage. Transition probabilities between disease stages (i.e., probability of progression) for patients with malignant nodules (local→ regional→distant) were informed by Gould et al., 2003,[109] (which also informed Edelsberg et al., 2018.[106]). The probability of progression is constant across disease stages for undiagnosed nodules. The paper does not report how the probability of progression from diagnosed regional to diagnosed distant disease was informed. Disease progression can occur beyond two years for some malignant nodules in Sutton et al., 2020, [107] whereas Edelsberg et al., 2018, [106] only explicitly models progression during the two years of CT surveillance.

The link to longer term-outcomes mortality outcomes is established by modelling transition to cancer specific death states, with the mortality risk conditional on whether the malignant nodule has been diagnosed (and therefore, treated) or not, the disease stage for diagnosed nodules and time on health state for diagnosed cancers. The mortality risk for undiagnosed malignant tumours appears to be independent of disease stage and constant in time, except potentially for distant cancers (not specified if diagnosed or undiagnosed) where the risk reduces over the first 4 years. Mortality risk for treated malignant nodules was informed by survival data from SEER data on NSCLC patients for local (T1N0M0, one small local nodule), regional (any T N1-3 M0, some regional nodes without metastasis) and distant lung cancer (any T any N M1, metastatic disease), taken from Gould et al, 2003.[109] Patients with diagnosed malignant nodules with local disease who survive for 5 years and those with diagnosed regional disease who survive and do not progress for 5 years transition to a disease-free state, so this model explicitly assumes cure for these patients. Patients with diagnosed distant disease have a lifetime stage specific mortality risk. The authors do not state what is the mortality risk in the disease-free state or diagnosed benign states, but this appears to correspond to age-adjusted all-cause mortality form UK life tables which is said to apply to all health states.

Cancer stage specific health-state utility values for malignant tumours was also sourced from the same source as for Edelsberg et al., 2018, although the estimates do not perfectly match between studies or with the source data. Age-adjusted utility values from the UK EQ-5D population norms are reported in the paper, but it is not clear if this apply just to patients with benign tumours of if any adjustment is made in malignant health states to reflect ageing of the population.

Costs in the model are linked to treatment, and do not depend on cancer stage. Treatment is assumed to include surgery alone or in combination with either chemotherapy or radiotherapy. A proportion of patients is assumed to have complications from surgical treatment, which have associated costs. All unit costs are sourced from NHS reference costs. No long-term disease or palliative care costs are considered in the model.

4.5 Conclusions of cost-effectiveness review of EarlyCDT Lung studies

There is limited evidence on the cost-effectiveness of EarlyCDT Lung on the diagnostic pathway for pulmonary nodules, with neither of the two studies identified being considered suitable to inform the current decision problem due to important differences in, for example, the patient population, the position and use of EarlyCDT Lung within the diagnostic pathway and exclusion of relevant diagnostic comparators, and the diagnostic accuracy evidence used to inform it.

The existing evaluations consider and quantify a number of components of clinical and economic value for EarlyCDT Lung in patients otherwise referred to CT surveillance, including: i) the increased cost of testing with EarlyCDT, ii) the cost and adverse events trade-offs of replacing CT surveillance by further investigations in those testing positive to EarlyCDT Lung, iii) the early detection of lung cancer (and potential stage shift) in the true positives to the test, and iv) the potential for increased detection of lung cancer if some of the true positives would have been missed by CT surveillance. The mechanism of value from increased detection is also expressed as early detection and assumes that cancers missed by CT surveillance would present clinically later in time. Despite overtreatment of indolent malignant nodules being of unclear relevance, neither model reflected the potential for increased overtreatment with the introduction of EarlyCDT Lung.

The evidence used to inform the population, the diagnostic outcomes, and the health and cost outcomes is sparse in many key aspects that will drive value such as the prevalence of malignancy and disease progression under CT surveillance. Modelling relies on unclear structural assumptions, without the support of relevant evidence. Therefore, the EAG considers that the evidence supporting the modelled effect on stage distribution (stage shift) is very limited.

The use of EarlyCDT Lung as part of a screening strategy for lung cancer has been evaluated in a large trial conducted in Scotland[13]. For the reasons described in Section 3.1.7.1, there is little relevance of this evidence to inform the clinical effectiveness of EarlyCDT Lung in the diagnostic pathway. As a consequence, the within-trial cost-effectiveness evidence is also of little relevance to this assessment.

Given the limited evidence on the cost-effectiveness of EarlyCDT Lung in the diagnostic pathway for pulmonary nodules, and to allow a fuller critical assessment of the assumptions and data sources used in the existing cost-effectiveness studies and to assist in the conceptualisation of a new decision model, further targeted literature searches for cost-effectiveness studies were undertaken. The review of the identified studies is reported in the next section.

5 ADDITIONAL TARGETED REVIEWS TO SUPPORT MODEL CONCEPTUALISATION

To support model conceptualisation, two further literature reviews of cost-effectiveness modelling studies were conducted: one on diagnostic tests or strategies within in the diagnostic pathway for pulmonary nodules, and the other on screening strategies for lung cancer. These technologies/strategies are expected to show common components of value to EarlyCDT Lung. Screening occurs upstream from diagnosis of lung cancer and, in common with the existing EarlyCDT Lung cost-effectiveness studies, cost-effectiveness models on screening use a mechanism for evidence linkage (based on stage-shift). It is hence important to consider this broader evidence as part of the conceptualisation and development of the new decision model.

Here we review the assumptions and evidence underlying such quantifications to inform the conceptualisation of a future assessment for EarlyCDT Lung.

5.1 Searches and studies identified

The searches retrieved 615 records of which 546 were excluded on the basis of title and or abstract. Full text publications were retrieved for 77 records and these were screened for potential inclusion on the reviews of diagnostic (28 titles) or screening studies (49 titles). The full text publications of two records identified at the first stage of screening for potential inclusion in the screening review were not retrievable, and were, therefore, excluded from the review.

Forty-five studies met the inclusion/exclusion for inclusion in the reviews. Ten of these studies were cost-effectiveness studies of diagnostic tests[106, 107, 109, 112-118]. Since two of the studies[106, 107] had already been reviewed in Section 4, only the remaining 8 studies were included in the review of diagnostic studies. Of the 36 screening studies[94, 119-125 126-152];[153] , one study the studies [120] did not report sufficient information to characterise the evidence linkage, and was excluded from the screening review. The remaining 35 studies[94, 119, 121-125];[126-153] were included in the screening reviews.

Details on both the diagnostics and screening reviews are reported in Appendix 11.8. A summary of the reviews is presented over the next sections.

5.2 Summary of the review of cost-effectiveness studies on diagnostics for lung cancer diagnosis

Table 23 identifies and briefly summarises the 8 diagnostics studies[109, 112-118] in terms of the population and important features of the sequences of diagnostic tests considered. It also summarises the three key components of each evaluation: the final classification, i.e., how the nodules were classified at the end

of the diagnostic strategy, the treatment choices (determined by the classification), and whether long-term health outcomes were linked to disease staging.

Table 23 Overview of diagnostic studies

Study(year), country	Population	Features of the test sequences considered	Classification	Choice component	Survival linkage via disease staging (Y/N)
D'Andrea (2020), US	Former or current smokers (screening population) with an indeterminate SPN	PET-CT vs. introduction of BGC in the test pathway either for central lesions only or for all lesions Possibility of referral to surveillance: Yes	(+) or (-)	(+): Surgery (fixed proportion of wedge resection, lobectomy and segmentectomy) (-): CT surv or discharge	Yes
Deppen (2014), US	Patients with pulmonary nodules (1.5- 2 cm) detected by CT and indication for suspected lung cancer without a preoperative diagnosis	Diagnostic surgery (VATS) vs. PET-CT vs. biopsy (CT-FNA) vs. bronchoscopy (NB) Possibility of referral to surveillance: Yes	(+) or (-)	(+): lobectomy (-): wedge resection or CT surv (leading to discharge)	Yes
Dietlein (2000), Germany	People with a SPN (≤3 cm) diagnosed by CT without calcification, spicula or enlargement of mediastinal lymph nodes	exploratory surgery vs. surv vs. biopsy (CT guided TNB) vs. PET Possibility of referral to surveillance: Yes	. benign lesion, locally resectable or unresectable cancer . for PET: with or without lymph node involvement	(+) resectable: surgery (+) unresectable: palliative care (+) with lymph node involvement: radiation (-): CT surv (leading to discharge)	Yes
Goehler (2014), US	Patients in whom pulmonary nodules were incidentally detected during CCTA (for CAD evaluation)	Surveillance vs. no follow-up Possibility of referral to surveillance: Yes	(+) or (-)	(+): lobectomy (-): CT surv (leading to discharge) or discharge	Yes
Gould (2003), US	Adult patients with a new, noncalcified SPN on chest radiograph	40 sequences of five diagnostic interventions: CT, PET-CT, biopsy, surgery and X-ray surv Possibility of referral to surveillance: Yes	(+) or (-)	(+): surgery (-): CT surv (leading to discharge) or discharge	Yes
Jiang (2020), US	Hypothetical population presenting with nodules at screening for CAD	CTCS vs. FCT Possibility of referral to surveillance: Yes	(+) or (-)	NR	Yes
Lejeune (2005), France	Incidental indeterminate SPN identified by standard chest X-ray.	Surv vs. PET vs. CT+PET Possibility of referral to surveillance: Yes	(+) or (-)	(+): lobectomy (-): wedge resection, CT surv (leading to discharge) or discharge	Yes

Rickets (2020), UK	Indeterminate peripheral SPN in which image-guided biopsy is recommended	EBN vs. TTA Possibility of referral to surveillance: Not explicit	(+) or (-)	NR	Yes

BGC, bronchial-airway gene-expression classifier; CAD, coronary artery disease; CCTA, coronary computed CT; CTCS, conventional computed tomographic calcium scoring; CT-FNA, CT guided fine-needle aspiration; EBN, electromagnetic navigation bronchoscopy; FCT, Upper lung field in addition a calcium scoring test to image the "full chest"; NB, Computer-assisted navigation bronchoscopy; NR, not reported; PET, positron emission tomography; surv, surveillance; TNB, transthoracic needle biopsy; TTNA, Transthoracic needle aspiration; VATS, video-assisted thoracoscopic surgery.

The study populations are diverse in terms of route of identification and positioning of patients in the diagnostic pathway. The two studies[115, 116]on patients with incidentally detected SPNs are both on patients undergoing investigations on the coronary artery disease diagnostic pathway. There is also significant variation in the strategies evaluated. All studies considered surveillance either as a strategy on its own or as part of the diagnostic pathway, with the exception of one study[118] (which compares bronchoscopy with needle biopsy and simply imposes a delay on false negatives). All studies considered a dichotomous classification (+ [malignant], - [benign]), except one study[114], which distinguished cancer according to its resectability and also considered the presence of lymph node involvement.

All studies appear to condition long-term health outcomes on disease staging.

The indirect value components (i.e. those relating to classification) identified across studies (see detail by study in Appendix 11.8, Table 37) were:

- Earlier diagnosis/ increased detection of lung cancer;
- Management of false positives with unnecessary follow-up tests and/or treatment, and decisions to treat benign nodules;
- Regression of benign nodules leading to early discharge from CT surveillance.

All the diagnostic studies modelled earlier diagnosis[109, 112-118], and all but one[113] considered increased detection. The increased detection compared to surveillance was imposed variably and relied mostly on assumptions on the specificity of CT surveillance or the uptake of CT surveillance rather than robust evidence (see Appendix 11.8.1.2, Table 38). As in the cost-effectiveness studies on EarlyCDT Lung (Section 4.4), both earlier diagnosis and increased detection were modelled via stage shift.

The delay to diagnosis with CT surveillance was modelled either by assuming that diagnosis occurred at a single specific point in time in the future or across multiple future time points (see details in Appendix 11.8.1.2); this was informed either by assumptions or by explicit modelling of nodule growth. The evidence used to inform models of nodule growth was not robust or appropriate (e.g., one study[109] used the same VDT that was used to inform the EarlyCDT Lung cost-effectiveness studies[106, 107], see Section 4.4.3). Another study modelled nodule growth and disease progression using an existing natural history model developed to simulate the outcomes of patients identified by screening[115], but insufficient detail is provided to characterise the evidence linkage and its appropriateness.

The delay to diagnosis was linked to disease staging by either assuming fixed stage shift for tumours with non-immediate diagnosis (e.g., all tumours diagnosed by CT surveillance progress from stage 1

to stage 2) or using a preclinical (i.e., before diagnosis) progression model. The assumptions in models reflecting a fixed stage shift from the delay to diagnosis were not justified. The models which included a preclinical progression component were informed by i) lung cancer screening trial data, [154] ii) VDT data collected with pre-CT imaging technology,[110] or iii) elicited evidence from public health policies to promote early diagnosis of lung cancer.[155]

Overdiagnosis of indolent malignant nodules is not modelled in any of diagnostic studies. Some studies consider treatment for a proportion of benign nodules, those that show growth during CT surveillance (see Appendix 11.8.1.3), and reflect this on short-term mortality and morbidity in the health outcomes and costs considered in the models, but evidence supporting malignant growth rates for benign nodules is not robust. Additionally, some studies consider the possibility of nodules presenting a negative biopsy being referred to treatment, reflecting that biopsy results may have limited bearing on treatment decisions.

Handling of false positives is detailed in Appendix 11.8.1.3. False positives at the end of the overall diagnostic strategy are handled in the identified studies by applying to the patients who undergo unnecessary surgical treatment the procedural mortality, HRQoL loss and costs associated with surgery.

In order to establish the link to final outcomes, the models conditioned outcomes on disease status and to disease stage for patients with lung cancer (see Appendix 11.8.1.2.). Survival outcomes of lung cancer patients were also conditioned on age. One study[115] included a competing mortality risk for CAD, thus reflecting comorbidity in the study population which was composed of patients with incidentally detected SPNs who underwent investigations for CAD. HRQoL of patients with lung cancer was conditioned on staging, histology, cancer recurrence of cancer, type of treatment and response, and time post-treatment. Only one study[118] conditioned costs of lung cancer patients on staging; other studies seemed to reflect mostly the costs of immediate cancer treatment with surgery. The health outcomes of patients with benign nodules were conditioned on age and sex, and generally reflect those of the general population. The models assumed that these patients did not accrue costs beyond those determined by the diagnostic pathway (procedural costs with or without complications).

Some studies considered the possibility of a proportion of benign nodules regressing during CT surveillance, but do not provide detail on how this component of value was modelled (see Appendix 11.8.1.3). Regression of benign nodules, may lead to early discharge from surveillance of a proportion of patients who will no longer incur the costs of CT surveillance and potentially assuage anxiety due to surveillance.

The review identified a single UK study [118], which used UK relevant evidence on long term survival, costs and HRQoL, all by disease stage at diagnosis. The study sourced other cause mortality from UK lifetables.

5.2.1 Key conclusions of the review of cost-effectiveness studies on other diagnostics for lung cancer diagnosis

Diagnostic studies use a stage shift mechanism of value that is consistent with the EarlyCDT Lung studies. These studies show that there is little or no empirical evidence supporting key aspects of model structure and key model parameters, particularly relating to quantifications of the delay to diagnosis with CT surveillance and associated stage shift. Also, the limited reporting of model inputs and results precludes assessments of validity. For example, the assumed speed of preclinical progression, important in determining the extent of stage shift, is only reported in one study.[118]

Across these studies, a number of additional components of value have been quantified variably. These include the possibility (or not) of benign resection and the possibility of differential detection across diagnostic strategies.

5.3 Summary of the review of cost-effectiveness studies on screening for lung cancer

As stated in Section 5.1, 34 studies on the cost-effectiveness of screening for lung cancer were identified by the searches. Given that the aim of the review was to have a general (but not comprehensive) understanding of how value components relevant to EarlyCDT Lung were modelled in the screening literature and the high volume of studies identified, we selected a sample of publications for review. This sample of screening studies aimed to include a sufficient range of modelling approaches. We also included in this sample all identified UK model-based cost-effectiveness studies, as the evidence used in these studies is more likely to be relevant to the UK context. The fully reviewed studies are identified and briefly summarised in Table 24 in terms of the type screening strategies considered (no screening vs. one-off screening and/or repeat screening), key features of the disease model, including the modelling approach, the sources of effectiveness data and whether survival outcomes were linked to disease staging.

Table 24 Overview of screening models

Study (year). Where there are multiple studies using the same model structure, differences are highlighted		Screening strategies			Disease model			
		No	One-off	Repeat	Modelling approach	Health states/ Staging	Main source of effectiveness data on early diagnosis/stage shift	Survival conditional on staging? (Y/N)
Snowsill (2018), Griffin (2020)	Two publications of the same model	Y	N	Y	Discrete Event Simulation	IA, IB, IIA, IIB, IIIA, IIIB, IV Cancer death, Other cause death	NLST	Y
Marshall (2000)	Different screening strategies evaluated	Y	Y	N	Decision tree, cohort	I, II, IIIA, IIIB, IV	ELCAP	Y
Marshall (2001)		Y	N	Y				
Yang (2017)		Y	Y	N	Mathematical model, cohort	I, II, IIIA, IIIB, IV by histology (SMLC, SqCC, non-SqCC)	NLST Scenario: NELSON + UKLS	Y
Pyenson (2012)	Pyeson (2012) and Peyson (2014) model different perspectives, and Vilanti considers HRQoL outcomes in addition	Y	N	Y	Cohort (actuary) model	A, B, C (assumed equivalent to local, regional, distant)	ELCAP for screened (NLST in a scenario) SEER for unscreened	Y
Pyenson (2014)								
Vilanti (2013)								
Ten Haaf (2017)	Different jurisdictions	Y	N	Y	Microsimulation	IA, IB, II, IIIA, IIIB, IV by histology (adenocarcinoma or large cell carcinoma or BAC; SqCC; other NSCLC, and SMCL)	NLST+PLCO (SEER also used in calibration)	Y
Tomonaga (2017)		Y	N	Y			NLST+PLCO (Swiss mortality statistics also used in calibration)	
Toumazis (2017)		Y	N	Y	Microsimulation	Early or advanced-stage, by histology (NSCL, SCLC)	NLST+PLCO	Y
Whynes (2008)		Y	Y	N	Decision Tree, cohort	NA	No stage shift	N
Field (2016), Field (2016a)	The studies by Field et al. are two publications of the	Y	Y	N	Decision Tree, simulation	I, II, III, III, IV	UKLS+UK cancer statistics	Y

Hinde (2018)	same model, and Hinde modifies the input evidence to reflect the Manchester lung cancer screening pilot				Manchester lung cancer screening pilot+ UK cancer statistics	
Hofer (2018)		Y N Y	MM. cohort	I, II, IIIa, IIIb, IV, no lung cancer, death	German Centre for Cancer Registry data (incidence)	Y

BAC, Bronchioloalveolar carcinoma; DT, decision tree; MM, Markov model; MS, microsimulation; SMCL, small-cell lung cancer; SqCC, Squamous-cell carcinoma

There is one key common mechanism by which screening strategies derive value compared to no screening, and this relates to earlier diagnosis arising from identification of pre-clinical cancer that would have otherwise only been clinically detected (this is commonly denominated 'lead time' in the screening literature). The link between early diagnosis and outcomes is mediated via a disease stage shift in almost all models. This is similar to the mechanism modelled in the diagnostic studies reviewed in Sections 4 and 5.2. The studies differ in terms of the clinical evidence used to inform the lead time estimates and stage shift and how this evidence is used. For example, some studies used (experimental) comparative effectiveness evidence of lung cancer screening LDCT to infer preclinical to clinical progression. [140, 141, 144, 146, 156]. One model estimated the probabilities of preclinical to clinical progression using cancer registry data.[129] Other models did not model preclinical to clinical progression, and used clinical effectiveness evidence differently. For example, some studies directly applied non-randomised evidence for stage distributions of screened vs. clinically detected lung cancer combined with assumptions on lead time and survival conditional on stage [16, 94, 125, 128, 137, 138, 149]

All studies except one [150] condition survival outcomes on stage at detection.

The value components relating to classification identified across studies (see detail by study in Appendix 11.8.2.2, Table 40) were:

- Earlier diagnosis (increased) detection of lung cancer;
- Earlier recalls resulting in some patients undergoing additional screening scans after a suspect result and incurring delays to diagnosis;
- Overdiagnosis of malignant indolent tumours;
- Management of false positives with unnecessary follow-up tests and/or treatment;
- Radiation exposure with increased cancer risk.

The studies established the evidence linkage required to model early diagnosis in screening models (see Appendix 11.8.2.3) in two main ways: i) by modelling preclinical to clinical progression or b) by linking effectiveness data on stage distribution combined with assumptions on lead time, to survival outcomes.

Where disease progression is explicitly modelled (see Appendix 11.8.2.3), the lead time and stage shift for screened vs. unscreened patients with lung cancer is quantified by tracking patients flow in the natural history model until detection (clinical or via screening). Overdiagnosis, i.e., the proportion of tumours that are detected with screening in excess of those clinically presenting with a no screening strategy is also a model output. The probabilities of progression from preclinical to clinical progression transition probabilities component were inferred using calibration methods and (mostly) comparative evidence from RCTs on lung cancer screening (e.g., NLST and PLCO in Ten Haaf et al.,

2017[145]). Preclinical to clinical progression probabilities are stage specific in these models, and two models further conditioned these probabilities on tumour histology.[144-146] All assume that pre-clinical progression is sequential across disease stages. One study [146] explicitly modelled the relation between tumour size, tumour growth, and metastatic spread, and linked it to disease progression (and probability of cure). Most of these studies do not model disease progression after lung cancer detection, the exception is Hofer et al., 2018[129], which models progression across three stages of 'after care' and further treatments (chemotherapy + radiotherapy or palliative treatment).

The models without a preclinical to clinical progression component (see Appendix 11.8.2.3) rely more heavily on assumptions and are more likely to be affected by bias. For example, two studies[94, 125, 128] used evidence on stage distribution for screened patients from screening studies (UKLS or Manchester lung screening pilot) but for unscreened patients from national cancer statistics – this implicitly assumes comparability between lung cancer patients participating in screening pilots and those clinically detected. Another issue with these models is that they require assumptions to model lead time, and these assumptions are not always robustly supported by evidence (see Appendix 11.8.2.3). Failure to appropriately model lead time, risk biasing survival estimates, which may be overestimated for patients with screened detected cancers. Lead time bias arises from screening prolonging the interval between diagnosis and death, (even if early treatment had no effect on patient survival), as diagnosis occurs earlier with screening compared with clinical detection. Thus, it is important that estimated survival benefits do not unduly incorporate lead time. Handling of lead time bias in models without a preclinical to clinical progression component varied; either by a direct adjustment on survival estimates (relying on assumptions) or a differences-in-differences methodology was applied to age adjust survival differences between screened and unscreened patients with lung (see Appendix 11.8.2.3).

Although models with a preclinical to clinical progression component do not rely solely on assumptions to estimate lead time, lead time bias can still arise in these models if additional constraints are not placed on survival. For example, one of the UK based models[140, 141] imposed the same lung cancer survival in each disease stage regardless of the type of detection (screening vs. clinical).

The survival of lung cancer patients (see Appendix 11.8.2.3, Table 42) was conditioned across most models on staging, histology and age. Some studies also conditioned the survival of these patients on detection type. One study [146] explicitly links survival to the probability of cure, which is conditional on tumour size and metastatic burden.

A common assumption across studies which modelled preclinical to clinical progression was that of no or negligible lung cancer mortality in preclinical stages (i.e., patients could only die of other

causes). One study [140, 141] explicitly allowed for early diagnosis within the same disease stage (comparing screening with no screening), so that for a proportion of patients there was no stage shift with early diagnosis. However, the model did not assume any survival benefit for early diagnosis in the absence of a stage shift, because the authors considered that evidence suggesting improved survival for screen-detected cancers vs. non-screen-detected cancer (when detected at the same stage) was at high risk of bias.

The HRQoL of patients with lung cancer was conditioned across models on staging, histology, detection type (clinical or screening) and histology, treatment and/or treatment type, time post-successful treatment, post-detection/treatment (clinical) health state, end-of-life, age, sex; the majority of studies conditioned HRQoL on staging, age and sex (see Appendix 11.8.2.3, Table 42). HRQoL was assumed to be constant over time (post-detection) or time varying i) with age or ii) assuming general population utility after 5 years disease free. One study[140, 141] assumed a temporary disutility from screening for both individuals with or without lung cancer to reflect anxiety associated with undergoing the intervention.

The costs of patients with lung cancer (see Appendix 11.8.2.3, Table 42) were also conditioned on staging across a number of studies. Costs were either assumed to be constant over time or time varying according dependent on time elapsed post-diagnosis/treatment and/or phase of treatment (initial vs. later treatment).

In the majority of models, the survival and HRQoL of individuals without lung cancer was conditioned on age/birth year and sex, with some models further adjusting estimates to reflect the characteristics of the population eligible for screening in terms of smoking status, exposure or history. The costs of individuals without lung cancer are not included in any of the models (other than the costs of screening and any further investigations.

Overdiagnosed lung cancers (see Appendix 11.8.2.4) in models with a preclinical to clinical component, appear to have the same outcomes of other true positives. Only one study explicitly states that constraints were placed on survival (e.g., the survival of stage specific of all lung cancers did not vary between screen and clinically detected tumours) to mitigate overdiagnosis (and other) bias(es). In models without a preclinical to clinical progression overdiagnosis was handled in scenario analyses where the survival benefit across the overall screened population was assumed to be smaller or by assuming an adjustment to prevalence with impact on costs and survival of overdiagnosed tumours. None of these scenario analyses were informed by evidence on the proportion of overdiagnosed tumours or their outcomes (see Appendix 11.8.2.4).

The majority of studies modelled the impact of false positive results to screening on outcomes as additional costs due to further unnecessary investigations (see Appendix 11.8.2.4). Only two models[129, 140, 141] explicitly linked false positives to survival to reflect the disutility associated with subsequent diagnostic follow-up and another to the associated mortality.[146]

Two components of value not considered in the diagnostic studies, but modelled in the screening studies relate to i) early recalls and ii) radiation exposure (see Appendix 11.8.2.4). One study[129] considered early recall CT scans for a proportion of patients who screened positive instead of proceeding directly to the diagnostic pathway. This was modelled as an additional costs and not linked to a delay to diagnosis. As mentioned above, Yang et al., 2017,[152] applied a lifetime cost to reflect the impact of radiation exposure due to screening to on patients who die from radiation induced cancer. It is unclear to whom this impact applies and how did radiation exposure differ across strategies.

Two sources of bias associated with early diagnosis, namely lead time and length bias, were considered in some screening studies but not in diagnostic studies. Screening models handled lead time bias in three ways (see Appendix 11.8.2):

i. constraining stage specific survival of patients with screen-detected cancers, so it did not exceed that of patients with clinically detected cancers;
ii. reducing the survival benefit of patients with screen-detected by an arbitrary amount of survival tie (not supported by evidence); or
iii. applying a differences-in-differences methodology to age adjust the survival differences between screened and unscreened patients with lung cancer.

Length bias was only explicitly discussed considered in one model.[140, 141] It was handled in the same way as lead time bias, i.e., by constraining stage specific survival of patients with screen-detected cancers, so it did not exceed that of patients with clinically detected cancers. It is worth noting that, as length bias arises from slow growing tumours being more likely to be detected by screening (given the interval between screening appointments; see Appendix 11.8.2.3), length bias also relates to overdiagnosis of indolent malignant tumours (an extreme case of slow growth).

A few studies (see Appendix 11.8.2.3) use UK relevant data sources to inform survival and costs by cancer stage. No UK specific HRQoL evidence was used to inform the outcomes of patients with lung cancer. UK relevant lifetables were used to estimate the survival of individuals without cancer. Survival and HRQoL adjustments to reflect the outcomes of smokers were also informed by UK relevant data.

5.3.1 Key conclusions of the review of cost-effectiveness studies of lung cancer screening

This review showed that the key mechanism of value attributed to screening in cost-effectiveness studies is of stage shift arising from earlier detection of lung cancer. This is consistent with the mechanism of value used in cost-effectiveness evaluations of diagnostics (including EarlyCDT Lung). Most screening studies evaluate screening in relation to clinical presentation (no screening). In this context, screening has been shown to lead to meaningful gains in terms of time to detection. Howvere, such level of gains in time to detection are unlikely to be observed with the use of EarlyCDT Lung in the diagnostic pathway where it may displace a CT surveillance strategy.

However, some screening models use the more robust clinical effectiveness evidence on screening to evaluate time to pre-clinical stage progression, a crucial quantity in linking earlier diagnosis to stage shift. In the absence of directly relevant evidence on the level of stage shift possible within the diagnostic pathway, a future assessment could consider the relevance of this evidence on pre-clinical progression arising from screening models.

A strength of the clinical effectiveness evidence on screening is that it is often grounded in high quality comparative studies on the stage distributions observed with earlier diagnosis (achieved via screening) or with a later diagnosis (at clinical presentation, and/or from different screening schedules). However, pre-clinical progression is, by definition, an unobserved quantity. Inferences over this are therefore established by calibrating pre-clinical progression models to multiple sources of observed data (including, but not solely, the abovementioned comparative studies). The robustness such calibration analyses is unclear because: i) the use of calibration makes it difficult to establish the contribution of different evidence sources, ii) reporting of the pre-clinical progression estimates is often poor, iii) sensitivity to alternative estimation assumptions is often not determined, and iv) despite a number of screening RCTs existing, there has been no attempt to consider this evidence together. The recent ECLS trial [13](see Section 3.1.7.1) could be included in the broader body of evidence informing speed of pre-clinical progression.

A number of other value drivers/components were quantified in these studies that could be relevant for EarlyCDT Lung. Some of these studies hypothesise that within-stage shifts may be associated with

survival benefits, despite none having quantified such an effect. Some of these studies consider the possibility of benign resection, from the imperfect specificity of the current diagnostic pathway of identified nodules. Additionally, some of the studies that use an evidence linkage approach to evaluate long term impacts on outcomes of stage shift take into consideration the potential for lead and length time bias. Finally, the potential consequences of increased radiation exposure could also be relevant.

5.4 Conclusions of the additional reviews

The additional reviews highlight that cost-effectiveness evaluations conducted within the diagnostic pathway for solid pulmonary nodules are generally based on sparse evidence. Despite the lack of evidence, these studies rely on a common (assumed) value mechanism: that diagnostic technologies displacing CT surveillance may lead to diagnosis of lung cancer at an earlier stage. Screening cost-effectiveness studies also use such value mechanism. The reviews identified a number of additional value components that could be of relevance for EarlyCDT Lung, for example, the potential for increased detection (i.e. the potential for the introduction of EarlyCDT Lung leading to a higher number of lung cancers detected). Finally, these broader reviews have helped identify structural assumptions and parameters estimates that could be used in alternative to those implemented in the EarlyCDT Lung cost-effectiveness studies. Many important gaps, however, still remain. These will be further systematised and explored in the following section.

6 CONCEPTUALISATION OF THE DECISION MODEL AND IDENTIFICATION OF EVIDENCE REQUIREMENTS FOR FUTURE ASSESSMENTS

This section identifies key considerations for the design of a decision model to support an assessment of EarlyCDT Lung (model conceptualisation), grounded on key evidence gaps and likely evidence requirements. It draws on the findings from sections 3 to 5 and on the judgements and views of the clinical expert that supported the EAG.

As the reviews in section 3 to 5 illustrate, evidence on EarlyCDT Lung and the diagnosis of pulmonary nodules is sparse, not only on the technology itself but also on the population of interest (e.g. prevalence of malignancy), the flow of patients, the clinical efficacy of the current diagnostic pathway, and the link between early diagnosis and long-term outcomes. Existing decision models are based on a number of assumptions that are unsupported by evidence, such as the extent of stage shift from avoiding referrals to CT surveillance. Faced with such uncertainty, the published cost-effectiveness analyses could have been accompanied by comprehensive and meaningful sensitivity analyses and value of information analyses, but none of the models reviewed does so to an appropriate extent. This limits the relevance of the conclusions reached. The EAG considers that the current analyses are not sufficiently robust to inform decision making.

In face of the evidential uncertainty, instead of aiming to identify a single model structure and recommend a particular modelling approach, the EAG outlines the key evidence requirements and main considerations for modelling, based on the value components identified in the suite of reviews conducted within this DAR (Sections 3 to 5). We will use influence diagrams (explained in detail in the following section) to identify the possible structural relationships needed for evidence linkage, and support future conceptualisation efforts that will be necessary as evidence on key aspects of the evaluation emerges.

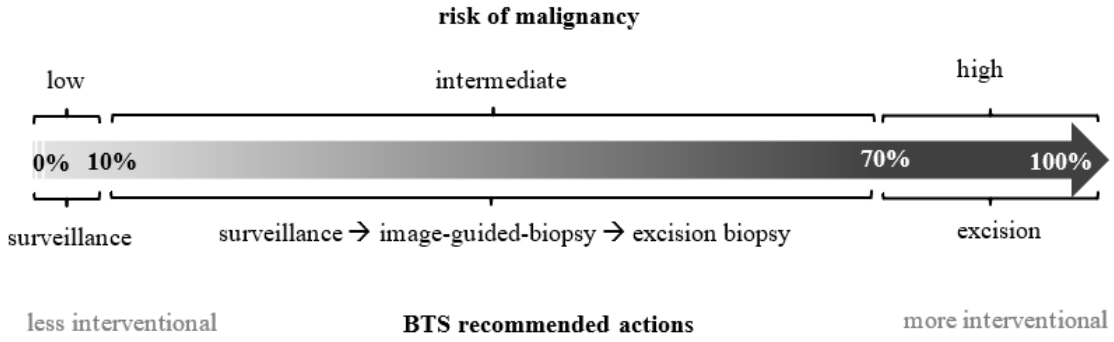
6.1 Core components of the decision problem

For most diagnostic technologies, such as EarlyCDT Lung, patient and health system benefit arises from the information the test provides which is used to tailor subsequent patient management decisions; value is therefore accrued indirectly.

In the context of this assessment, EarlyCDT Lung is being considered to be included in the diagnostic pathway for solid pulmonary nodules. The BTS pathway (Section 1.1.2), commonly used in the UK, grounds management decisions, which range from CT surveillance (less interventional) to excision (more interventional), on numerical assessments of malignancy risk (Figure 1). EarlyCDT Lung test results are being proposed to update these malignancy risk scores.

Clinical decisions are explicitly grounded on two risk thresholds: one determining referral to CT surveillance (<10% risk) and another referral to excision (>70%) (Figure 17). The guidelines are less prescriptive for the intermediate risk group, recommending image-guided biopsy but also allowing the use of CT surveillance and excisional biopsy. Clinical decisions for this risk group are determined on a case by case basis and depend on risk of malignancy, considering the net trade-offs of further interventions for individual patients (including the patient's fitness to undergo invasive diagnostic follow-up and subsequent treatment), patient preference and nodule characteristics (e.g. nodule location, where peripheral nodules will be easier to access than central ones).

Figure 17 BTS recommended actions according to malignancy risk



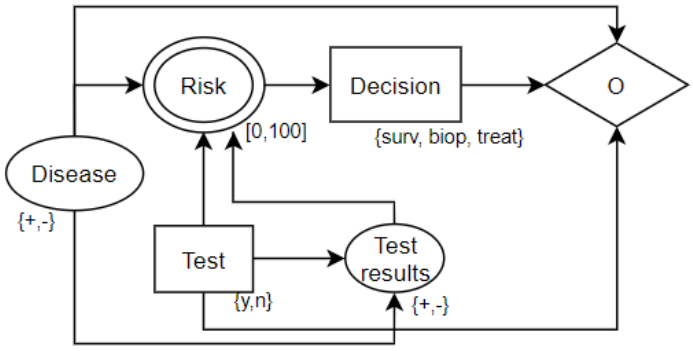
To support conceptualisation of a future decision model, and to illustrate some of the considerations arising in subsequent sections, we will use influence diagrams,[21, 22] which provides a simplified representation of the decision problem. These diagrams use shapes to represent important aspects of the evaluation – rectangles represent deterministic events (such as decisions), ovals represent probabilistic events (events that are uncertain) and diamonds represent the outputs of interest. Arrows between shapes reflect dependencies, which only matter if they directly or indirectly affect outcomes.

The influence diagram in Figure 18 represents the core components of the decision problem for EarlyCDT Lung. In the diagram, disease status (Disease) is represented as a chance node, reflecting the probability of malignant (+) or benign (-) disease. The malignancy risk score (Risk) is probabilistic (represented by a distribution), and because it is a continuous variable (between 0 and 100%) the shape is represented using a double line. The arrow from Disease to Risk indicates that the risk score is determined by malignancy status, i.e., the risk score distribution is expected to differ between benign and malignant nodules. Options for management decisions within the BTS pathway (Decision) are surveillance (surv), biopsy (biop) or treatment (treat), and these are determined by the risk score. The diagram represents treatments as deterministic decisions from risk scores. This means that for a given risk score, a single decision is taken (in later sections this assumption is relaxed). The

risk score here is shown as continuous, but the score could also be categorised (e.g. 0-10, 10-20, 20-50, 50-70 and 70-100 as done in Section 3.3.2.3) to simplify the representation of how risk scores determine management decision. These management options, alongside disease status, will impact on outcomes (O), an output of the model. This would include both the short-term impacts of the management decisions, but also long-term effects of treating malignant nodules.

A decision node is used to reflect the decision to include a new test (Test), such as EarlyCDT Lung, in the diagnostic pathway. The direct arrow from Test to Risk illustrates the case where the test is not used. When the test is used, its results (Test results) update the quantitative risk score. The diagram reflects that the Test is assumed to affect further management decisions only by changing the risk score. The test can itself have direct impact on outcomes (represented by the arrow between Test and O), reflecting its costs and any adverse events.

Figure 18 Influence diagram – Core components of the decision problem



This core conceptualisation diagram identifies important aspects of this evaluation, which will be looked at in further detail in the next sections. These include:

- Population, particularly in what concerns value drivers such as prevalence of disease (Section 6.2),
- Subsequent clinical management decisions and how EarlyCDT Lung affects these (Section 6.3), and
- How changes in subsequent clinical decisions affect outcomes (Section 6.4).

6.2 Population

In this section we summarise the evidence available on the characteristics of the populations and subpopulations of interest (described in full in sections 1.2 and 1.5, and listed in Table 25 below), and highlight important issues around subsequent actions determined by test results, which are fundamental in determining the clinical and economic value of EarlyCDT Lung.

Evidence on the population with pulmonary nodules is sparse, of unclear representativeness and is heterogeneous. [3] This includes evidence on characteristics that drive value for a new diagnostic test such as prevalence of disease (as shown in Edelsberg et al., 2018[106]- see section 4.3). This is reflected in existing cost-effectiveness studies, where value drivers have been informed by either evidence of limited relevance (e.g., the use of Tanner et al., [108] to inform prevalence in Sutton et al.,[107] as critiqued in Section 4.4.1), or unsubstantiated assumptions (such as stage distribution [Sections 4.4.3 and 5.2]). A single small UK study by Al-Ameri[74] described the flow of patients through the BTS pathway (described in Section 3.2.6). This study suggests that more than half of patients with incidentally detected nodules present small or low risk nodules with a low prevalence of malignancy, and that approximately one third present with intermediate risk and a higher prevalence of malignancy. A non-negligible proportion of cancers detected at metastatic disease were observed across both risk groups.

6.2.1 Evidence required and modelling considerations

The Al-Ameri study[74] represents the best evidence on the UK population on which to base an economic model. However, it is of small size and therefore future evidence collection efforts should focus on describing the (sub)populations of interest, including the size of the population and key characteristics that drive value, such as prevalence, diagnostic or surveillance procedures used, histology and stage distribution at diagnosis. Given the differences between the subpopulations in the prevalence of malignancy highlighted by the Al-Ameri study,[74] future cost-effectiveness studies of EarlyCDT Lung should establish value separately for each subpopulation.

It is important that future evidence helps understand and describe potential sources of heterogeneity. For example, two cost-effectiveness models of diagnostics focussed on nodules incidentally detected in patients undergoing workup for coronary artery disease (Goehler et al.[115] and Jiang et al.[116], Table 23), suggesting that the reason for CT scan is a potential source of heterogeneity. More broadly, characterisation of heterogeneity across patients (e.g., emphysema, route of presentation) and nodule characteristics (e.g. size, location), would be valuable, particularly as some of these characteristics may be associated with malignancy risk, speed of nodule growth, speed of pre-clinical progression and/or long-term health outcomes.

6.3 Clinical decisions under current pathway and clinical impact of EarlyCDT Lung

Clinical evidence on EarlyCDT Lung that would be required for an economic model is discussed in Section 3.4.4 and includes

a. robust diagnostic accuracy evidence on the population and subpopulations of interest,
b. validation of pre- and post-test risk scores and
c. evidence on clinical impact of EarlyCDT Lung in changing subsequent management decisions.

Section 6.1 identified that important impact on patient outcomes from the use of EarlyCDT Lung arise from the changes in management it can lead to. The range of possible actions after risk assessment with EarlyCDT Lung are listed in section 1.5.1. The evidence reviewed in Section 3 and clinical advice indicated a number of further relevant considerations:

- Management decisions in the intermediate risk are heterogeneous, with the proportions referred to CT surveillance, biopsy or excision being largely unknown.
- Some nodules are difficult to biopsy, such as sub-centimetre nodules and nodules centrally located in the lung. This restricts management options to either CT surveillance or excision.
- The value of EarlyCDT Lung in determining malignancy risk is unclear. The EAG analysis (Section 3.1.5) found poor diagnostic accuracy of EarlyCDT Lung, and consequently, based on EAG modelling (Section 3.3), a limited impact on risk of malignancy. For example, an individual with a pre-test risk of 10% would obtain a maximum post-test risk score of 22%, and a post-test risk of 70% can only be achieved in individuals with a pre-test risk above 48% (see Figure 10).
- The widespread availability of PET-CT means that all patients in the UK are expected to have access to this technology. Where Brock risk is reclassified to above 10% after EarlyCDT Lung, patients are expected to receive PET-CT to inform further management decisions.

Based on these considerations, the potential for changes in management in the proposed positionings for EarlyCDT Lung are (further detail presented in Table 25 below):

- EarlyCDT Lung is unlikely to change referrals to CT surveillance in a number of subgroups, including small and low risk nodules that cannot be biopsied. EarlyCDT Lung is therefore unlikely to present clinical or economic value in these groups.
- CT surveillance → biopsy: low or intermediate risk nodules that would have been referred to CT surveillance but that can be biopsied. Note that the intermediate risk nodules considered here are likely to show a lower pre-test risk score (close to 10%). At the range of 10-48% pre-test risk, EarlyCDT Lung cannot lead to post-test risk scores above 70% (under the EAG's analyses, Table 17), therefore it is unlikely that these nodules will see their management change from CT surveillance to excision;
- CT surveillance → treatment: intermediate risk nodules with a pre-test risk score above 48% and that cannot be biopsied; and
- Biopsy → treatment: intermediate risk nodules with a pre-test risk score above 48% and that would have been biopsied.

Table 25 Management under current practice and with the addition of EarlyCDT Lung

Subpopulation	Current management	Possible management choices for those with increased post-test risk after EarlyCDT Lung		
		CT surveillance	Biopsy	Excision

1	**small nodules**	CT surveillance	Y	N*	N**
2	**low risk nodules**	CT surveillance	Y	Y if eligible for biopsy	N**
3	**intermediate risk nodules**	If not eligible for biopsy: . CT surveillance	Y	N	Y, if pre-test risk score is sufficiently high
		If eligible for biopsy . CT surveillance (likely to present a lower pre-test risk)	Y	Y	N***
		. Biopsy	N	Y	Y, if pre-test risk score is sufficiently high

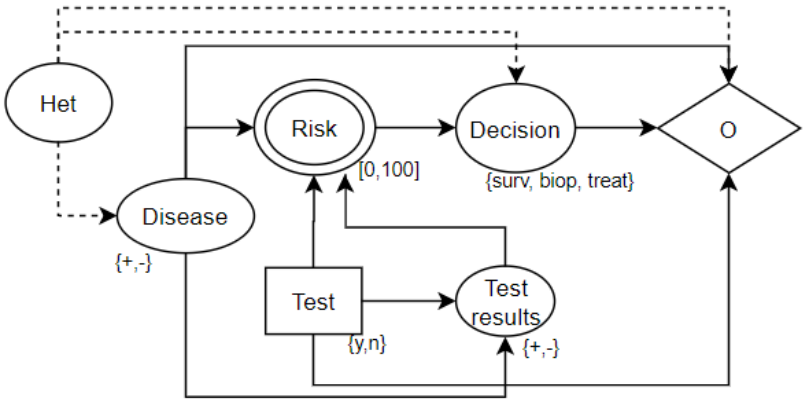
* sub-centimetre nodules cannot be biopsied, ** nodules with a pre-test risk <10% cannot see their post-test risk increased to above 70%; *** under the EAG's model, EarlyCDT Lung cannot return post-test risk scores above 70% in nodules with a pre-test risk below 48%

6.3.1 Evidence required and modelling considerations

It is important that further research allows a better understanding of how the Brock or Herder risk of malignancy scores are used to inform clinical management decisions. In recognising that there is variability in management decisions, particularly in the 10-70% range, future evidence should explore the relationship between risk of malignancy and the likelihood of referral to surveillance and excision. Additionally, evidence discerning how factors such as patient preference and fitness to receive more invasive tests contribute to these decisions is currently unavailable. It is important to consider the potential impact of such variation in clinical practice in decision modelling to accurately predict outcomes and obtain unbiased results from the economic modelling. The influence diagram in Figure 19 modifies the diagram in Figure 18 to include a probabilistic relationship between risk score and management decisions: management options are no longer represented by a rectangular (deterministic) node as in the previous diagram (Figure 18), but by an oval chance node reflecting that, for each value of the risk score, there is a probability of referral to surveillance, biopsy or excision.

The EAG's analysis in Section 3 shows that the extent to which EarlyCDT Lung leads to changes in management depends on the test's accuracy. Further evidence emerging on the accuracy of EarlyCDT Lung should therefore be carefully considered in future modelling attempts, and interpreted in the context of the test's ability to affect subsequent management choices. Direct evidence on how EarlyCDT Lung test results affect subsequent management decisions would also be important to support assumptions over its clinical utility.

Figure 19 Influence diagram – expanded diagram to reflect expected variation in management decisions and to consider heterogeneity



In this section we have listed important considerations on management decisions relating to the different subpopulations/positionings for the test, and future assessments should explicitly consider these. Note that none of the published cost-effectiveness studies on EarlyCDT Lung have considered these (Section 4). Future modelling efforts should reflect subgroups with restricted vs. unrestricted management options (e.g., D'Andrea et al. [112] restricted options for diagnostic follow-up with biopsy of central nodules to only those with a diagnostic bronchoscopy result), which will include people with nodules that can, or cannot, be biopsied, and people at higher/lower risk of serious adverse events from biopsy.

Section 6.2 identifies important sources relating to patient and nodule characteristics which are linked to prevalence of disease. The reasons determining variation in management decisions considered here may also be related to prevalence of disease, particularly those related to nodule characteristics (e.g. small nodules). Therefore, in the influence diagram in Figure 19, heterogeneity (Het) is broadly considered. The diagram illustrates that sources of heterogeneity can determine prevalence of disease (arrow from Het to Disease) and subsequent management decisions (arrow from Het to Decision). It also represents the possibility of sources of heterogeneity affecting outcomes directly (arrow from Het to O), which is also be important to be considered in further decision modelling (e.g., histology of malignant tumours, Section 6.4.2.2).

6.4 Components of clinical and economic value for EarlyCDT Lung arising from changes in management decisions

In this section, we focus on the link between changes in subsequent management decisions arising from EarlyCDT Lung's clinical utility in the diagnostic pathway for solid pulmonary nodules and outcomes. These highlight key trade-offs (components of value), arising indirectly via changes in management decisions, that are relevant to consider against the cost of introducing the test itself and

any adverse events or anxiety introduced by the test (which affect all individuals tested, and have been previously detailed in Sections 1.3.1.1 and 3.1.7.1) when determining the clinical and economic value of EarlyCDT Lung. These have been identified by bringing together the issues/limitations from the different reviews (sections 3-5) and are:

- The short-term impacts (costs and adverse events) of escalating the current pathway to more interventional diagnostic investigations/treatments on positives to the test for which management is changed. These include: i) the costs and harms imposed by unnecessary invasive diagnostics or treatments on benign nodules (false positives) and indolent nodules (true positives that would not have shown significant growth on CT surveillance) and ii) the implications of radiation exposure from increased referral to PET-CT scan.
- Longer term health benefits and cost implications of earlier detection (and treatment) of lung cancer in true positives to the test for which management is changed, and/or increased detection from the overall diagnostic strategy that includes the test (i.e. a higher proportion of true positives in relation to current pathway).

These key components are further linked to the clinical utility of EarlyCDT Lung in Table 26, highlighting that the trade-offs arise as a consequences of changes in management: short-term impacts arise on both true and false positive patients that see management change (facing the risk of overtreatment of benign and indolent nodules and the potential for increased radiation exposure), and that the long terms effects will only be realised for the true positives that see management escalated, leading to early or increased detection (and consequent treatment) of malignant lung cancer.

Table 26 Components of value of EarlyCDT Lung arising from changes in further management decisions

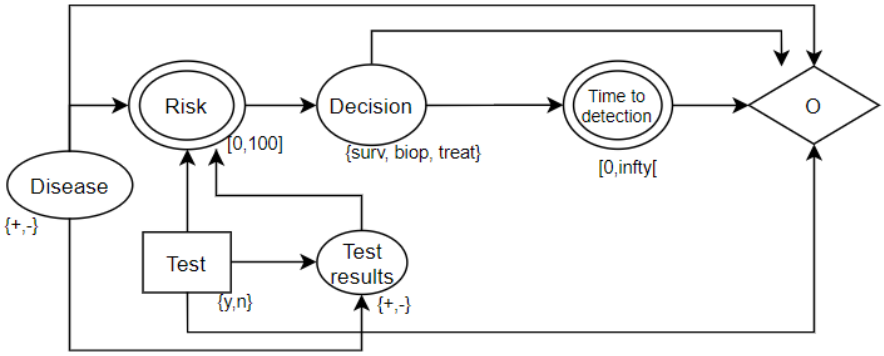
Components of value	True positives for which management changes	False positives for which management changes	All negatives and any positives with unchanged management
Short term impacts of replacing current strategy with further diagnostic investigations and treatments	Impact of escalated diagnostic/treatments, including intervention on indolent nodules	Impact of escalated diagnostic/treatments, including unnecessary intervention on benign nodules	--
Health benefits and disease cost reductions from increased detection and/or earlier detection of clinically significant cancer	. increased detection, if current strategy has imperfect sensitivity . earlier detection, if strategies differ in the time to diagnosis (e.g. surveillance)	--	--

Evidence required and modelling considerations

Future cost-effectiveness models for EarlyCDT Lung should clearly justify the trade-offs quantified and include consideration for each of the value components in Table 26. The influence diagram in Figure 20, expanded from Figure 19, highlights the two key components of value from changes in management. First, the direct effects of these choices over outcomes (represented in Figure 20 by the direct arrow from Decision to O) which includes their costs and adverse events. Note that, for example, a surveillance strategy (current management for many of the subpopulations here considered) will include further diagnostic workup in malignant nodules showing quick growth. Therefore, net impacts from escalation will arise only from higher detection of clinically significant and from differences between the strategies in detecting of indolent disease.

The second component of value reflects the longer-term health benefits and cost savings arising indirectly from earlier detection (and treatment) of malignant disease (represented by the addition of the event 'Time to detection', that links management decisions to outcomes).

Figure 20 Influence diagram – expanded diagram to reflect key components of value for EarlyCDT Lung arising from changes in management decisions



The following subsections summarise existing evidence, and identify further evidence requirements and the evidence linkages necessary to support economic modelling on these two components: section 6.4.1 focusses on short-term impacts of escalating subsequent diagnostic/treatment decisions and section 6.4.2 focusses on the longer-term impacts from increased/earlier detection of lung cancer. The balance of each of the components of value will differ for each of the proposed placements for EarlyCDT Lung; this is discussed in the concluding subsection (Section 6.5).

6.4.1 Short-term impacts of escalating diagnostic/treatment

The short-term impacts of the escalation in management relate to costs and adverse events, and include the level of unnecessary intervention (and ultimately, of benign resection). These will depend on the likely shifts in management from the introduction of EarlyCDT Lung (see Section 6.3). Next,

we present considerations on future quantifications of these trade-offs, which are also summarised in Table 27.

6.4.1.1 CT surveillance

Section 6.3 identifies that EarlyCDT Lung is likely to lead to displacement of CT surveillance for more interventional procedures in two of its proposed positionings (low and intermediate risk nodules). Section 3.2.6, however, identified limited evidence on the clinical impacts of undergoing CT surveillance within the BTS pathway. This is also reflected in the models reviewed in Sections 4 and 5.1 which are largely underpinned by assumptions.

Implications for the modelling of longer term impacts of increased/early detection are further detailed in section 6.4.2 ahead (covering uncertainties in the sensitivity and extent of delay to diagnosis imposed by CT surveillance). Of relevance for the cost-effectiveness modelling of short-term impacts of CT surveillance are the level of referral to further unnecessary diagnostics/treatments. This is associated with the false positive rate, itself determined by the specificity of CT surveillance (Section 3.2.6) and by the prevalence of malignancy (which will vary across the subpopulations and is therefore important to be explicitly modelled by subpopulation). With regards to the costs of CT surveillance, it is important to determine the mean number of scans until either referral to further diagnostics or discharge (including early discharge due to nodules disappearing at subsequent scans). Evidence is required on the probability of referral/discharge at the different scan points established in the BTS guidelines. Mean number of scans can also be formally modelled using evidence on VDT measurements; however, the EAG did not identify any existing source that was robust and contemporary (Sections 4 and 5), and would therefore recommend further evidence collection. No significant adverse events are expected from CT surveillance.

Decision criteria for CT surveillance are based on nodules presenting significant growth, and therefore, indolent (but malignant) nodules may be less likely to be identified with surveillance than with other diagnostic strategies. The overdiagnosis of indolent lesions (that are unlikely to cause harm) is often cited as a concern in the early diagnosis of cancer, particularly in screening studies. Indolence is, however, typically associated with subsolid lesions on CT but has been documented in solid lesions[157]and therefore cannot be clinically ruled out. One of the EarlyCDT Lung cost-effectiveness studies[106] considered 18% of overdiagnosis of malignant nodules (based on data from a lung cancer screening population); the rate of overdiagnosis was, however, assumed common between the EarlyCDT Lung strategy and the CT surveillance strategy.

The extent of indolent disease in solid nodules is largely unknown, and further evidence on its prevalence and the likelihood of overdiagnosis under CT surveillance and under alternative diagnostics is therefore required.

6.4.1.2 Biopsy

EarlyCDT Lung may affect the likelihood of patients receiving biopsy in two of its proposed positionings. The first positioning includes low risk patients classified by the Brock score that see their post-test risk increased above 10% following EarlyCDT Lung and subsequent PET-CT scan and are therefore diverted from CT surveillance to biopsy (non-surgical) or bronchoscopy. The second positioning includes intermediate risk patients that would have otherwise received biopsy or bronchoscopy but may be referred to direct excision by EarlyCDT Lung. For economic modelling, evidence requirements on biopsy/bronchoscopy procedures to evaluate these two positionings include:

1. evidence on how post-test risk score and clinical management could change in the first group of patients, particularly after re-evaluation with the Herder score after PET-CT imaging. Additionally, there should be consideration for the potential for increased radiation exposure with PET-CT, which is higher than with CT surveillance. Evidence on its consequences is therefore required to support decision making.
2. the breakdown, in clinical practice, between the use of biopsy and bronchoscopy (noting that their indications for use do not entirely overlap, and that the availability of augmented bronchoscopy is limited);
3. the accuracy of these two procedures which determines the rate of benign resections (together with the impact of test results in decisions about excision, see item 5 below). Current evidence (see Section 3.2.7) establishes that biopsy presents a higher overall accuracy (noting that this being significantly reduced in small lesions due to increased diagnostic failure and lower sensitivity) then bronchoscopy;
4. risk of complications, which current evidence (Section 3.2.7) establishes is elevated with biopsy, such as pneumothorax (the risk of which is determined by lower FEV1 and presence of emphysema along the needle tract) bleeding and air-embolism;
5. acknowledge that, due to the possibility of false negatives, negative results to biopsy/bronchoscopy may have limited bearing in management decisions or lead to the procedure being repeated. Variation in how negative biopsies determine repeat biopsy and management decisions would need to be explicitly considered in a future assessment; and
6. of particular relevance to the second positioning here considered, is to determine whether pre-surgical biopsy/bronchoscopy adds delay to treatment in relation to direct excision, and the implications of such delays to the outcomes from surgery.

Table 27 : Considerations on short-term impacts of escalating diagnostic/treatment

	Description	Considerations on the value components	Considerations on costs	Potential AEs of relevance	Other considerations
CT surveillance strategy	Low dose CT scans at multiple timepoints (complex schedule), followed by further diagnostics/treatments for patients with nodules showing growth.	. specificity determines %s getting unnecessary further diagnostics/treatments	. consider how patients flow through the surveillance schedule (including discharge and further referral), to determine the average number CT scans . consider costs of further diagnostic/treatments	. lower radiation exposure than PET-CT .anxiety from time under surveillance	. important to consider histology and prognosis according to VDT . sensitivity determines increased detection . how early will malignant nodules be detected determines delay to diagnosis
Non-imaging tests and non-surgical biopsy	Image-guided biopsy; augmented bronchoscopy	. consider eligibility for bronchoscopy and CT-guided biopsy and how delay in diagnosis may affect eligibility.	. % of bronchoscopy vs biopsy . consideration for the % of non-diagnostic samples in biopsy . consideration for the need for repeat biopsy where a negative result is obtained	. pneumothorax , bleeding and air embolism, which occur with higher incidence in biopsy	. biopsy can better guide excision . little value in low or high-risk patients as management options are unlikely to change
Surgical and non-surgical treatment	Surgical: VATS or thoracotomy; wedge, lobectomy or segmentectomy. Non-surgical: SABR or RFA	. consider the need for explicitly linking primary tumour treatment to outcomes, which could allow reflecting within-stage gains . explicitly model benign resection and its consequences	. breakdown of treatment modalities across disease stages at diagnosis, and consider potential for within stage differences . costs categories should include treatment costs, and complications	. mortality . morbidity (e.g. respiratory complication, prolonged hospital stay, sepsis)	

A future assessment will also need to consider any new technique developments for biopsy and bronchoscopy, which may improve safety and accuracy (particularly for smaller nodules assessment). Additionally, it is worth noting that while, currently, pre-surgical biopsy is only required where it may influence treatment, the emergence of adjuvant or neoadjuvant treatments may mean that a pre-treatment biopsy specimen becomes always required.

6.4.1.3 Primary tumour treatment

The majority of diagnostic and screening models reviewed in Sections 4 and 5 do not explicitly model the link between primary tumour treatment and long-term outcomes (the exception being Hofer et al. [129]). Instead, treatment is implicitly embedded in the outcome data conditional on stage of disease at detection (considered in further detail in Section 6.4.2.2 ahead). The validity of this approach relies on assuming that differences in treatment modality and outcomes can be fully explained by disease stage. However, in the context of earlier detection, two factors may justify different outcomes of treatment of nodules which would be smaller but potentially still within the same stage of disease. The first is that surgical treatment requires accurate identification of lesion localisation, which may be more difficult in smaller lesions. The second is that less invasive primary tumour treatments, such as segmentectomy or even ablation, may be preferred in smaller nodules. Therefore, a future assessment needs to carefully consider the need and value of explicitly modelling primary tumour treatment, and any additional requirements this may impose in terms of evidence linkage to longer term outcomes.

The costs, morbidity and post-operative mortality impacts of alternative primary treatment options for early lung cancer are, however, often considered in decision models (Section 5), to distinguish primary tumour treatment impacts across the different disease stages at detection. To do so, it is important to understand the different treatment modalities used in clinical practice, which should include: the use of non-surgical treatment, the use of pathological confirmation at wedge resection, the use of VATS vs. thoracotomy and the use of lobectomy vs. anatomical segmentectomy. There is uncertainty about the current level of use of the different treatment modalities across disease stages. The risks of morbidity and mortality are significant (90-day mortality for lobectomy is estimated at 4%) [3] and vary across the modalities used, but the magnitude of differences in complications and oncological outcomes is uncertain. Beyond the treatment costs themselves it may be important to consider differences in waiting times for surgery, and in post-operative length of stay and total hospital costs. Future assessments should also consider that clinical practice may increase adoption of anatomical segmentectomy (due to its lower rate of complications) if evidence arises on how to better target this to patients.[158]

Evaluations, particularly for positionings of EarlyCDT Lung where resection without pre-operative confirmation of malignancy is considered, should explicitly consider the rate and consequences of

benign resection. There is uncertainty about the current level of benign resection, with rates reported in the literature as low as 2% (UK screening studies[92, 93]) or as high as 86% in a case series of indeterminate pulmonary nodules undergoing surgical excision[3] .The rate of benign resection should depend on the prevalence of disease in each of the subpopulations of interest (and the subgroup of patients within that may be brought forward to surgical treatment) and on the specificity of the overall diagnostic strategy used to support decisions to proceed to treatment. Additionally, it should depend on how decisions to treat are made, and in the variation in these decisions observed in clinical practice. These should depend on malignancy risk (determined by the BTS pathway for decisions on excision without preoperative confirmation of malignancy), the level of fitness for surgical treatment, and other factors, such as histology and stage of disease.

It is therefore important that further evidence is generated to provide a better understanding of the rate of benign resection. To allow explicitly considering how the level of benign resection may be affected by the introduction of EarlyCDT Lung in the different positionings, future modelling attempts should explicitly link the diagnostic accuracy of the overall diagnostic pathway and the prevalence of malignancy (across the subpopulations and/or subgroups of interest) to the level of benign resection (e.g. Deppen et al,[113] and Dietlien et al,[114] in Section 5.3 and Appendix 11.8.1.3).

6.4.2 Longer term impacts from increased/earlier detection of lung cancer

By facilitating earlier treatment, the earlier detection of lung cancer provides an opportunity for improvements in overall survival. Earlier detection has been demonstrated to be linked to detection at earlier stages of disease, i.e. to stage shifts, in randomised controlled trials of screening (in relation to no screening, i.e., clinical detection). Some of these RCTs have also shown reductions in lung cancer mortality. The most recent study demonstrating stage shift and mortality benefit is NELSON [95]. Stage shift and mortality benefits have been further modelled to establish the long term clinical and economic value of alternative screening strategies (Section 5.3).

To the EAG's knowledge, there is no experimental evidence of early detection or stage shift from alternative diagnostic strategies for incidentally detected nodules (Sections 4 and 5.2). However, this has been the key mechanism of value for EarlyCDT Lung in existing cost-effectiveness studies. It is therefore important that further research generates evidence to support this mechanism of value for diagnostic strategies in general, and EarlyCDT Lung in particular, ideally using an experimental design.

In the cost-effectiveness analyses reviewed, one of the models for EarlyCDT Lung and a couple of models on diagnostic strategies also assumed increased detection, attributing additional value to the

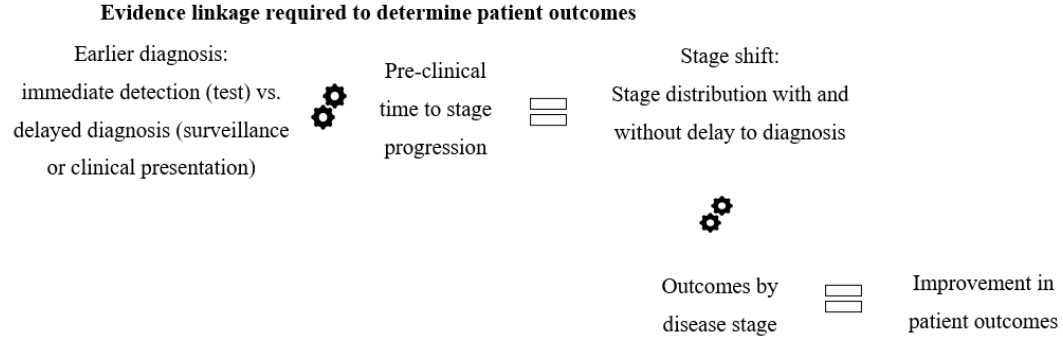
diagnostic technologies evaluated. Increased detection relies on assuming that comparator strategies (e.g., CT surveillance) fail to detected a proportion of cancers (usually small), which would present clinically at a later point in time. The new strategy, introducing an additional test with relevant sensitivity (such as EarlyCDT Lung), would therefore have the potential to detect these lung cancers earlier. The value of increased detection also lies on the earlier detection mechanism. However, there is no empirical evidence supporting this value component for EarlyCDT Lung.

The sensitivity of the overall CT surveillance strategy determines the potential for increased detection with EarlyCDT Lung. Clinically, it is considered that there is no growth rate threshold beneath which, nor duration of radiological stability beyond which, malignancy is definitely excluded.[3] There is, however, uncertainty concerning the proportion of clinically significant cancers missed by the BTS CT surveillance schedule. Further evidence on the likelihood of a malignant cancer being missed by surveillance is therefore required to support such an assumption.

To establish the value of early detection (including increased detection) in the absence of empirical evidence directly on the magnitude of stage shift attained, evidence linkage is required. The following mechanism (also illustrated schematically in Figure 21) has been used across the diagnostic and (most) screening studies reviewed in Sections 4 and 5, encompassing:

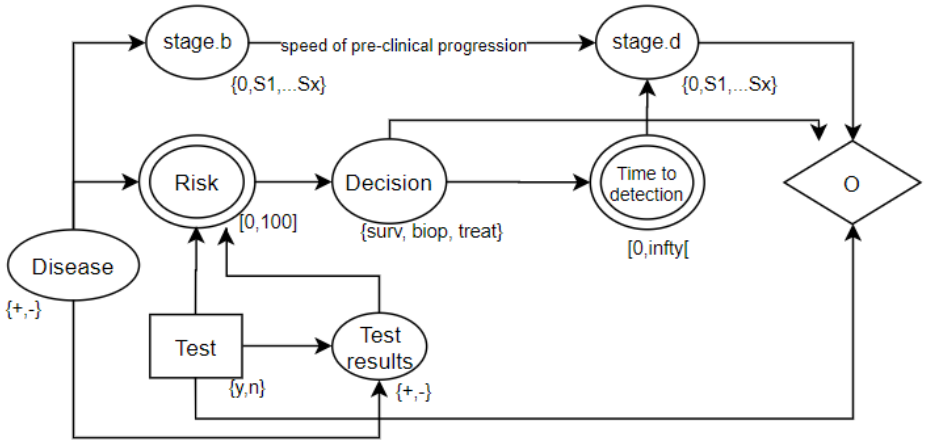
> i) the identification of differences in the time to diagnosis between current and proposed identification strategies, and mapping of these differences against likelihood or time to pre-clinical stage progression, to define the level of stage shift, and
>
> ii) the linking of the stage distributions, with and without stage shift, to expected long term outcomes conditional on disease stage.

Figure 21 Schematic representation of the evidence linkage required to establish value from early detection



One possible representation of such mechanism of value is shown in the influence diagram in Figure 22 below. In this diagram, the stage distribution at baseline (stage.b), i.e. at the time of first CT scan, is represented to include the absence of malignancy (using the value zero) alongside the categorisation in disease stages (S1 to Sx). Management decisions determine time to detection (which may assume values between zero, reflecting immediate detection, and infinity, reflecting no detection). This should be parameterised to consider that CT surveillance imposes a longer time to detection than biopsy or treatment; in this way, the shift from surveillance to biopsy/treatment that EarlyCDT Lung may facilitate will reduce time to detection. As explained above, by exposing time to detection, both i) the value of earlier detection from immediate diagnosis with biopsy or immediate treatment in relation to CT surveillance, and ii) the value of increased detection of cases that would not have been diagnosed with current care and would therefore only have presented clinically, much later, can be explored. Time to detection determines the stage distribution at detection (stage.d), with consideration for the speed of pre-clinical progression.

Figure 22 Influence diagram – expanded diagram to include the stage shift evidence linkage mechanism

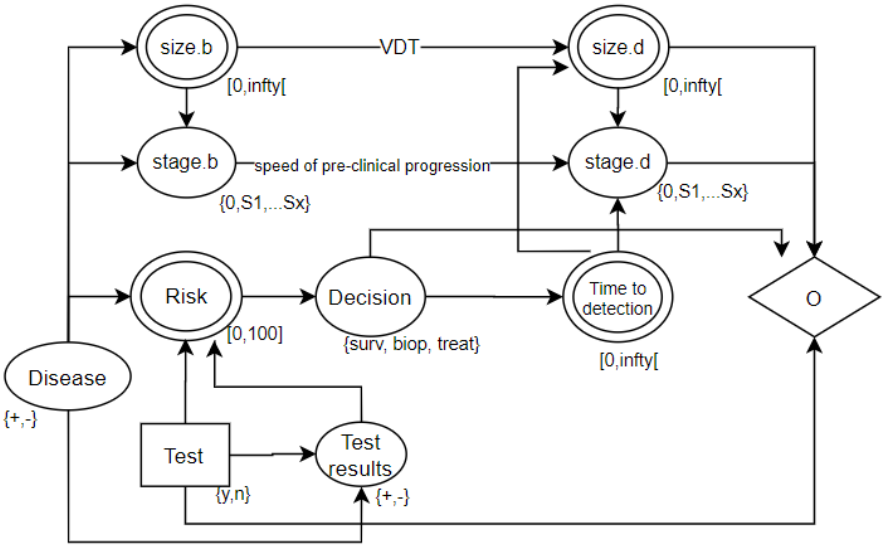


It is worth noting that disease staging classifications use discrete categories. However, there also needs to be some consideration over whether earlier detection within the same disease stage can also be associated with long term benefits. Stage classifications are mostly based on criteria such as size of nodule, location, lymph node involvement and metastatic spread. The latest TNM staging system emphasises the difference in prognosis between stages T1a and T1b that only differ in the size of the tumour. This implies a relationship between size of the tumour and health outcomes which suggest that earlier detection within the same stage (i.e. not allowing nodule growth) may be associated with improvements in the outcomes of treatment. Only one of the screening cost-effectiveness studies

reviewed explicitly modelled the relationship between growth and stage progression (section 5.3) but neither this study or others modelled within stage benefits, despite the potential for such an effect having been discussed. [140, 141] This is an area where further evidence is required. If evidence emerges supporting the benefits of within stage detection, the association between nodule size and health outcomes from treatment could be used, alongside metastatic burden, location and histology.

The influence diagram below (Figure 23) illustrates a possibility for tracking the size of the nodule at detection (using VDT and time to detection) in future decision models, to allow quantifying within-stage benefits. The tracking of the size of the nodule is represented in a similar way to the tracking of disease stage. The size of the nodule at detection is determined by the size of the nodule at baseline, VDT and time to detection. Given that size is one of the dimensions considered in most staging of disease classifications, in the influence diagram size of the nodule is represented to determine stage of disease. To allow within stage growth to affect outcomes, the diagram reflects that both stage and size determine outcomes.

Figure 23 Influence diagram – expanded diagram to reflect within-stage benefits (via tracking of nodule size)



Further details on evidence and modelling requirements in relation to this mechanism are discussed next.

6.4.2.1 Stage shift

Time to diagnosis: In the diagnostic pathway of detection of solid pulmonary nodules, the most significant source of delay is the possibility of referring patients to a surveillance strategy. Surveillance

refers to a schedule of regular imaging screens aimed at measuring nodule growth to identify VDT, as significant growth (below a certain threshold of VDT) is commonly associated with malignancy. Given the need for time to establish VDT, surveillance imposes a delay to diagnosis. Delay is determined by the probability of surveillance detecting clinically significant cancer at the different schedules screening points.

Section 3 highlights the absence of evidence on the overall sensitivity and timing of diagnoses with CT surveillance. The models identified in our review of diagnostic technologies/strategies (Section 5.2) either rely on unsubstantiated assumptions or on limited evidence of questionable relevance. Some of the models, including the EarlyCDT Lung cost-effectiveness studies (further details in Section 4 and Appendix 11.8.1.2) infer probability of detection from further modelling of tumour size and growth. The evidence underlying these VDT 'submodels' is limited, lacking relevance and robustness, and failing to characterise heterogeneity. Heterogeneity in VDT can be associated with patient and nodule characteristics, such as nodule size, probability of disease spread, histology and others [3]. This would need to be considered explicitly to appropriately determine probability of detection at different time points.

A key source of variation is histological subtype, which is likely to be related to size, VDT (progressively longer VDTs were identified for small cell carcinoma, squamous cell carcinoma, adenocarcinoma and bronchioalveolar carcinoma/adenocarcinoma *in situ* [3]) and outcomes. It is therefore important to reflect on whether the distribution of histologies detected may differ over the different timings of CT surveillance scans. This has not been explored in previous models (Sections 4 and 5), but could be considered in future modelling.

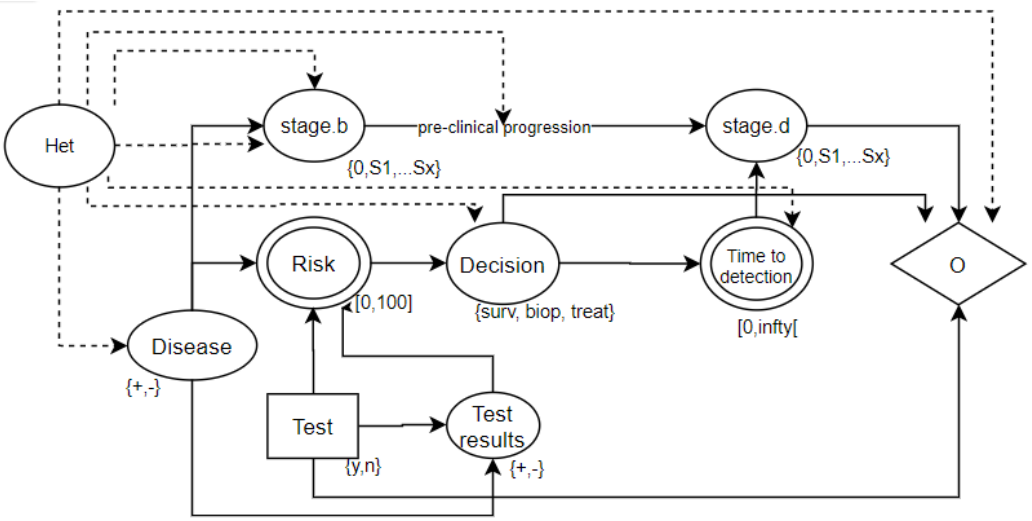
Stage progression: Despite tumour size being one of the features defining disease stage, with tumour growth therefore inherently determining progression, there is no evidence that stage progression happens within the timeframe of CT surveillance (see Section 3.2.6). Most of the models reviewed of diagnostic technologies, relied on unsubstantiated assumptions to define the likelihood of progression during surveillance (e.g., Dietlein et al.[114] and Lejeune et al., [117] detail provided in Section 5.2). Evidence on the likelihood of stage progression with CT surveillance is therefore required to support a future assessment of EarlyCDT Lung.

It is worth noting that the likelihood of stage progression should depend on the stage classification used (most diagnostic cost-effectiveness studies reviewed in Section 5.2 use 3-stage, local, regional, distant, or 4-stage, I – IV, classifications). The use of more granular categories, that is, of a more disaggregated level of staging categories (e.g. T1a distinct than T1b, or stage IA1 distinct from stage IA2), could allow stage-shift based evidence linkage approaches capture additional benefits that are currently not captured, reducing the impact of ignoring potential within-stage benefits.

Given the absence of evidence on the likelihood of stage progression for incidentally detected nodules followed-up by CT surveillance, wider evidence on the speed of pre-clinical stage progression is valuable. Screening RCTs provide a particularly robust foundation for evaluating the speed of pre-clinical stage progression. These studies typically compare clinical detection (no screening) with early detection from screening and observe the stage distributions at detection across the groups (which differ in time to detection). Further modelling uses these data to infer time to preclinical progression, based on the assumptions imposed in a natural history structural model and by calibrating such a model to a variety of data sources (examples are: Tomonaga et al.,[144] Ten Haaf, et. al., [145]; Toumazis et al.,[146] Griffin et al., [140], Snowsill et al., [141], and Hofer et al. [129]). The reporting is not sufficiently detailed to allow comparing estimates of mean time to progression across studies, and the influence of structural assumptions over these estimates (Appendix 11.8.2.3); further research on this would be welcomed.

Heterogeneity in stage progression is clinically acknowledged, and has been considered in a few of the cost-effectiveness models reviewed (see [146] as an example). The influence diagram below (Figure 24) exemplifies a number of different ways in which heterogeneity can affect aspects related to stage progression. The diagram already includes heterogeneity determining the prevalence of malignancy, restricting management options and determining outcomes, which have been previously discussed (Sections 6.2 and 6.3). Heterogeneity affecting disease progression is also exemplified in the diagram. For example, tumour histology may affect likelihood of progression (two of the screening models identified in Section 5.3 have conditioned pre-clinical progression probabilities on tumour histology[144-146]). Another case is where heterogeneity, for example in histological subtype, is associated with the stage distribution at baseline. Slow growing nodules will be more likely to be picked up incidentally (or by screening strategies), and faster growing nodules are more likely to be identified at an advanced stage of disease with potentially limited capacity to benefit from the level of earlier diagnosis expected from averting surveillance (this is associated with length time bias in screening studies). The final example presented reflects the possibility of heterogeneity, for example in nodule location, affecting the likelihood and time to detection. Further evidence should consider characterising heterogeneity in stage progression, linked to size (or growth), and could also include histology or other patient and nodule characteristics. Future decisions models should reflect such heterogeneity explicitly.

Figure 24 Influence diagram – expanded diagram reflecting the potential impact of heterogeneity



6.4.2.2 Long term outcomes

Outcomes of detected malignant nodules, by disease stage

As highlighted in previous sections, to link stage shift to health outcomes, the majority of the studies reviewed conditioned survival and HRQoL on disease stage at diagnosis (further detail in Sections 4 and 5). Other common health outcome determinants modelled in the screening and diagnostic studies include age, sex and tumour histology. It is important that future modelling carefully considers these and other possible determinants, identifying which may be more broadly relevant across the value mechanism proposed. For example, histology is related to health outcomes, but also to tumour growth and via this to the probability of the tumour being identified via CT surveillance. It is therefore important to reflect this source of heterogeneity across the entire evidence linkage mechanism.

Another consideration that arises when modelling survival outcomes of patients with malign tumours is that of competing mortality risks, as these risks may limit the ability to benefit from early diagnosis of patients. Some diagnostic studies in patients with pulmonary nodules incidentally identified in the context of coronary disease investigations have modelled the impact on mortality of CAD. As noted above (Section 6.2) there is considerable heterogeneity in the population defined by the decision problem and comorbidities (with impact on survival) are likely to vary by identification route. For example, patients identified via screening, usually aged 50 years or older and with high smoking exposure, may have cardiovascular and respiratory comorbidities with increased mortality risk compared to the general population. For patients whose nodules were identified incidentally the profile of competing mortality risks may also vary according to the reason for referral for the original CT scan. Patients referred to CT imaging after a trauma event are likely to be younger and with fewer

comorbidities than patients undergoing CT in cardiovascular diagnostic pathways. Since the proportion of patients with nodules identified via the different routes is uncertain, as are the patient characteristics within each subgroup, it will be challenging to accurately reflect the competing risks in the overall population. However, the decision model should be flexible enough to incorporate competing risks by population subgroup, so the impact can be explored in subgroup and/or sensitivity analyses.

Across the majority of the cost-effectiveness diagnostic[109, 112-118] and screening studies, [94, 125 137;128, 133, 134, 138, 140, 141, 149;144-146, 150, 152] the evidence linkage between stage shift and survival was established directly via disease stage without being mediated by treatment. The evidence used to inform survival outcomes is typically sourced from observational data such as cancer registries (e.g., SEER), which reported the survival outcomes of patients treated for lung cancer by disease stage at diagnosis (as well as age, sex, and histology). The use of registry data is usually driven by the need to have sufficiently long follow-up to capture impacts on mortality. However, the evidence does not reflect survival outcomes of patient treated with more contemporaneous lung cancer treatment (primary tumour treatment and subsequent treatments), but rather with the treatments available when the data was collected. In recent years, surgical techniques have advanced and a range of new treatments for lung cancer have become available, including a number of mutation-specific targeted therapies and immunotherapies.[159] Therefore, the use of registry data that may not reflect the outcomes and costs of patients treated for lung cancer over more recent years, introducing additional uncertainty concerning the magnitude of survival benefits linked to earlier diagnosis of lung cancer. It is worth noting that the effectiveness of newer lung cancer treatments will also have associated uncertainties, as the evidence will be less mature than for earlier treatments.

An alternative approach to establish the link between stage shift and survival would be to further conditional survival on treatment. Modelling the effect of lung cancer treatment could better characterise the survival of patients with earlier lung cancer diagnosis compared to current practice, but would present additional practical and evidential challenges. In order for the stage shift link to outcomes to also be mediated by treatment, additional evidence would be required, including:

i. Treatment allocation conditional on stage, histology and presence of treatment relevant biomarkers (e.g., anaplastic lymphoma kinase or programmed death-ligand 1)
ii. Distribution of treatment relevant biomarkers in the population (and its correlation with histology)
iii. Characterisation of subsequent treatment sequences and their health outcomes

The biases introduced in the evidence linkage necessary to quantify the impact on outcomes of the early diagnosis component of value, namely lead time and length bias, are highlighted in the screening studies (example in [140, 141]). The selection of an approach to handle lead time and length bias (three

alternatives have been identified in Section 5.3) in a future assessment should take into account the adequacy of the method to the model structure (e.g., does the model explicitly track disease progression from preclinical to clinical?), and make use of good quality evidence on survival gains with early detection of lung cancer.

Outcomes of undiagnosed lung cancer

In the majority of cost-effectiveness studies reviewed, the evidence linkage mechanism explicitly includes the diagnosis of lung cancer at clinical presentation if these are undetected by other means (incidental or through screening). Most models assume that lung cancer, while undiagnosed, has similar outcomes to the general population (i.e., its clinical significance is limited); examples that assume differential outcomes are Sutton et al,[107]and Hofer et al. [129]. A future assessment should consider evidence on the clinical significance of undiagnosed lung cancer.

Outcomes of benign nodules

The long term health outcomes of patients with benign nodules have been implicitly considered equivalent to those of the general population in previous diagnostics models, and individuals were assumed to not accrue costs beyond those determined by the diagnostic pathway (see Section 5.2). No robust evidence was identified to support this assumption. The prevalence of malignancy may differ across positionings for EarlyCDT Lung (and other potential factors such as route of presentation), and the rate of benign resection is expected to differ across strategies. Because of this, if there are differences in the longer term outcomes of benign nodules (such as those resulting from long term morbidity caused by benign resection), these should be explicitly considered in future modelling.

6.5 Conclusions

There is currently insufficient evidence to support an explicit quantification of the value of EarlyCDT Lung in the diagnostic pathway of solid nodules. Our reviews identified that, to justify the additional costs and health system implications of introducing EarlyCDT Lung in the BTS diagnostic pathway, the short-term trade-offs of escalating diagnostic/treatment (including overtreatment of indolent lesions and benign resection) should be considered against the long-term benefits that may arise from earlier identification of lung cancer. A number of important uncertainties arise, but based on current evidence and clinical judgement, it can be established that EarlyCDT Lung is unlikely to present value in small nodules (between 5 and 8mm), in low risk nodules that are not eligible for biopsy, and in intermediate risk nodules with a (pre-test) risk score below 48% that would undertake biopsy in the current pathway, as EarlyCDT Lung has limited ability to change management decisions in these groups.

Whether EarlyCDT Lung presents clinical and economic value for the remaining subpopulations will be determined by explicit assessments of:

-- in low risk nodules eligible for biopsy:

- The likelihood of EarlyCDT Lung changing management decisions (likely to be from surveillance to biopsy).
- The prevalence of malignancy (expected to be below 6%) and the accuracy of EarlyCDT Lung followed by biopsy, which will determine the probability of detection. This is to be compared to the accuracy and timing of detection with CT surveillance (and subsequent investigations) to determine the potential for early detection
- The stage distribution of the nodules at the time of initial identification (noting that a proportion may already be at advanced stages) and the likelihood of disease progression under surveillance, which determines the potential benefits of early diagnosis with EarlyCDT Lung
- The prevalence of malignancy and specificity of EarlyCDT Lung, which will determine the likelihood and consequences of escalating management in false positives to the test. Because pre-test risk is low and EarlyCDT Lung has limited ability to increase risk score, benign resection is unlikely

-- In intermediate risk nodules that would be assigned to CT surveillance in the current BTS pathway (these are, therefore, likely to be at the lower end of the risk spectrum) but on which biopsy can be undertaken:

- The likelihood of EarlyCDT Lung changing management decisions (likely to be from surveillance to biopsy). This is more likely than in the low risk population due to the higher pre-test risk score;
- The prevalence of malignancy is expected to be low (although higher than in the low risk population), therefore the net benefits of early detection may be low; and
- Given the low prevalence, the likelihood of increased intervention (biopsy) on benign nodules is of concern. As with the previous group, benign resection is unlikely.

-- In intermediate risk nodules presenting risk scores above 48%, that would be assigned to biopsy in the current BTS pathway:

- Likelihood that EarlyCDT Lung changes management decisions (likely to be from biopsy to excision).

- The potential for early detection may be limited, as surveillance is not on the current pathway for these nodules and time to treatment may not be not significantly changed. Potential benefits of in the true positives may only amount to avoiding biopsy: its cost, mortality and morbidity.
- The higher prevalence of malignancy in this group determines a lower likelihood of increased intervention of benign nodules; however, here intervention is likely to mean resection given that pre-operative confirmation of malignancy is not obtained. Resection has important morbidity, mortality and cost implications.

-- Intermediate risk nodules presenting risk scores above 48%, that would be assigned to CT surveillance in the current BTS pathway for not being eligible for biopsy:

- Higher likelihood that EarlyCDT Lung changes management decisions (likely to be from CT surveillance to excision).
- Given the higher prevalence in this group, there is a higher potential for early detection and stage shift.
- Due to the higher prevalence of malignancy, a lower likelihood of increased intervention of benign nodules is expected in this group, but this is likely to be benign resection

The potential for EarlyCDT Lung to lead to overtreatment of indolent lesions that would otherwise not be detected by surveillance is unclear, as the prevalence of solid slow growing nodules is unknown but likely to be very small. The potential for EarlyCDT Lung to lead to increased detection is also unclear. Clinically, the presence of malignancy is not ruled out after little or no growth being observed within a CT surveillance schedule; however, its probability is thought to be extremely low.

Considerations for a future assessment of EarlyCDT Lung

A future assessment of EarlyCDT Lung needs to ensure that the evidence supporting quantification of the abovementioned value components in the different groups is robust enough to support decision making. Table 28 below, summarises the evidence requirements (adapted from the NICE evidence standards framework [160]) and considerations for modelling for a future assessment of EarlyCDT Lung. Critical aspects are the prevalence of disease in each of the groups, the potential for harms of CT surveillance (in terms of delay to diagnosis and the likelihood of stage progression), and the clinical utility of EarlyCDT Lung in updating the risk scores commonly used to support management decisions.

To support evidence linkage approaches, it would be desirable to also have a better understanding of the sources and implications of heterogeneity in this patient population, particularly as some factors may affect the entire evidence linkage pathway, through to outcomes. For example, associations between histology and growth (VDT), would affect the likelihood of detection over the different time points of CT surveillance, i.e. resulting in different times to (delayed) detection of different histologies. The

association between histology and outcomes then determines the overall impact of health outcomes of these associations.

Table 28: Summary of evidence requirements and considerations for modelling for a future NICE assessment of EarlyCDT Lung

Key economic information	Evidence requirements	Considerations for modelling and evidence linkage
Population and subpopulations	.Description of the (sub)population(s) of interest on key drivers of value, including size of the group, mean age, sex and prevalence of malignancy, and, for those with malignant disease, description of stage distribution at initial identification and histological subtype. . Better understanding of heterogeneity over the prevalence of disease, particularly reflecting on factors that may be also linked to likelihood of detection/magnitude of stage shift and outcomes	. Explicit modelling of subgroups to reflect the different proposed positionings for EarlyCDT Lung . Consideration of relevant sources of heterogeneity, such as histological subtype, across the evidence-linkage mechanism.
Care pathway	. Flow of patients through the care pathway in the BTS guideline, and breakdown of clinical actions, particularly in the intermediate risk group. . Better understanding of how risk of malignancy, and other factors, determine subsequent management decisions	. Reflect variation in management decisions and how this is related to risk of malignancy. . Consider that variation may arise from personalisation of care (i.e. judgements over the balance of benefits and harms of more interventional procedures)
Effectiveness: accuracy	. Evidence obtained in a setting relevant to the UK health and social care system in the target (sub)population(s)/groups, demonstrating consistent benefit including in accuracy and in the validity of post-test risk scores. Potential sources of heterogeneity should be examined, e.g. patient and nodule characteristics. . A well-conducted meta-analysis if there are enough available studies on the technology.	. Consideration for the link between accuracy, post-test risk scores (and their validity), and the clinical utility of EarlyCDT Lung .
Effectiveness: clinical utility	. Comparative evidence (with a relevant comparator) on the clinical utility of the test in determining subsequent management decisions, with exploration of heterogeneity.	. Evidence on clinical utility could be directly included in the model, and/or integrated with accuracy and clinical utility information to explore generalisability of findings
Effectiveness: extent of earlier diagnosis and stage shift	. Comparative evidence (with a relevant comparator) on the extent of earlier diagnosis and stage shift, with appropriate consideration for potential heterogeneity.	. Evidence on stage shift could be directly included in the model and/or integrated with other sources within an evidence linkage approach to explore generalisability of findings
Long term health outcomes	. Evidence on the impact of early diagnosis on long term outcomes (within and across disease stages)	. Evidence linkage is likely required based on stage at detection. The use of disaggregated disease stage categorisations should be explored. The representativeness of sources of evidence on outcomes conditional on disease stage should be considered. The relevance of sources of heterogeneity should be considered.
Potential for escalation of interventions in benign nodules	. Evidence on the likelihood of benign nodules receiving non-surgical biopsy/bronchoscopy and resection (and the breakdown of surgical modalities received)	. Examine the relevance of benign resection for each positioning of EarlyCDT Lung using the evidence linkage approach.

Other value components	. Evidence demonstrating the applicability of other value components, including the potential for increased detection	. Explore the plausibility and relevance of including other value components in analyses.
Costs	. Cost parameters informed by costs relevant to the health and social care decision maker. Suitable sources include NHS reference costs or national tariffs. All costs associated with the interventions should be considered.	
Resource use	. Resource use parameters are based on study, pilot or real-world usage data, or on information obtained from relevant clinical or social care professionals or other appropriate sources. Show that the resource use parameters for the existing care pathway are validated as an accurate and comprehensive itemisation of resources currently used (including any variations by subgroup and over time) by evidencing approval and support from relevant professionals in the UK health and social care system. Show that the resource use parameters for the new care pathway are validated as an accurate and comprehensive itemisation of resources necessary and expected to be used in the new care pathway (including any variations by subgroup and over time) by evidencing approval and support from relevant professionals in the UK health and social care system.	
HRQoL	Health Related Quality of Life data measured using an appropriate standard measure, such as the EQ-5D. A rationale for the choice of measure should be provided. Show that the data have been collected in an appropriate way.	

A cost-effectiveness model supporting a future assessment should incorporate any emerging new evidence, including emerging mechanistic evidence which can be used to justify structural assumptions on the design of a future decision model. A future assessment should extensively and explicitly explore any remaining evidential and mechanistic uncertainties, and their impact over clinical and cost-effectiveness. The influence diagrams presented here can support future conceptualisation efforts and should be used as a basis for any further modifications which may be required to reflect emerging new evidence. These diagrams can also be used to define alternative assumptions where evidence remains less robust.

8 DISCUSSION

8.1 Statement of principal findings

8.1.1 Clinical effectiveness

The evidence on the use of EarlyCDT Lung specifically in people with pulmonary nodules is currently very limited. There are only 5 cohorts reporting 695 patients with nodules who have received EarlyCDT Lung, including 97 cancer cases. Only two of these cohorts have been fully published; the other three are only available as conference abstracts. In none of the cohorts was it explicit that EarlyCDT Lung had been received according to the proposed diagnostic pathway (see Figure 3).

Consequently, the existing evidence is at high risk of bias: most data on EarlyCDT Lung is not in people with pulmonary nodules, is outside of the proposed diagnostic pathway, or has issues regarding the timing of EarlyCDT Lung relative to identification of nodules or malignancy. This also means that the applicability of the existing evidence to the BTS diagnostic pathway is uncertain. The EAG notes that there has been very little investigation of EarlyCDT Lung without Oncimmune involvement. The EAG therefore considers that the existing evidence is insufficiently extensive and robust to be able to draw any firm conclusions on the diagnostic accuracy or clinical value of EarlyCDT Lung.

The evidence that does exist suggests a low diagnostic accuracy of EarlyCDT Lung. From bivariate meta-analysis, the EAG estimates a diagnostic accuracy for EarlyCDT Lung on its own of 20.2% sensitivity (95% CI 10.5 to 35.5) for a specificity of 92.2% (95% CI 86.2 to 95.8). This is notably poorer than estimates used by Oncimmune (e.g. 41.3% sensitivity at 90.6% specificity from Healy et al 2017).

Poor diagnostic accuracy may mean than EarlyCDT Lung can add little when combined with existing approaches in the diagnostic pathway, such as Brock or Herder risk assessment. EAG analysis of how using EarlyCDT Lung might alter pre-test risk found that having a positive EarlyCDT Lung test may only slightly increase the estimated risk of malignancy; for example, from 10% to 20%, or 50% to 70%. This means that it is unclear whether using EarlyCDT Lung would change clinical decision making for most patients.

The Brock risk model was found to have good diagnostic accuracy (AUC 92%, 95% CI 90 to 95, 8 cohorts), but data were too limited to assess diagnostic accuracy at key risk cut-offs, such as the 10% risk cut-off. The Herder risk model (after PET-CT) also had apparently good diagnostic accuracy (AUC 84%, 95% CI 77 to 92, 5 cohorts), although with limited data explicitly on Herder risk assessment, and no data sufficient to assess accuracy at key risk cut-offs. Given the apparent low diagnostic accuracy of

EarlyCDT Lung, and the higher accuracy of Brock and Herder risk assessment, this would suggest that adding EarlyCDT Lung to either test is unlikely to substantially improve diagnostic accuracy.

Although several meta-analyses of the use of PET-CT in patients with pulmonary nodules were identified, the studies included in these meta-analyses did not report the performance of PET-CT based on nodule size or on pre-test likelihood of malignancy, as categorised in clinical guidelines. Further searches identified only two studies which stratified results either by pre-test risk or by nodule size.

The EAG identified limited data on the diagnostic accuracy or clinical value of CT surveillance. One study found that using volume size and doubling time may have very high diagnostic accuracy to detect malignant nodules. It is therefore unclear whether using EarlyCDT Lung to move patients out of CT surveillance would offer clinical benefit.

There is adequate evidence providing diagnostic accuracy estimates methods for CT-guided transthoracic needle biopsy. Better quality studies of radial probe endobronchial ultrasound (r-EBUS)-guided transbronchial lung biopsy are needed.

The EAG identified no evidence on the clinical impact of using EarlyCDT Lung, such as how many patients would see a change in their diagnostic approach with a positive result. The EAG performed a simulation study to attempt to assess this, but limited data meant that the simulation rests on numerous assumptions, and may not be conclusive. The simulation study suggested that EarlyCDT Lung is unlikely to offer meaningful clinical improvement for low-risk nodules. At the 10% risk cut-off there was almost no difference in diagnostic accuracy between using Brock risk with EarlyCDT Lung versus using Brock risk alone. Consequently, the numbers of patients with malignant nodules who moved out of CT surveillance appeared to be small, and there would be rather more patients with benign nodules wrongly moved out of CT surveillance.

EarlyCDT Lung may have some use in identifying malignant nodules among those classified as intermediate risk (10% to 70%) after Herder risk assessment. Adding EarlyCDT Lung to Herder improved test sensitivity at the 70% risk cut-off. Patients with higher pre-test risk (e.g. above 50%) with a positive EarlyCDT Lung test would move to having a post-test risk of over 70%, and so might be considered for excision. These patients mostly had malignant nodules, with fewer false-positives. However, the risks of excision in the patients with benign nodules and a positive EarlyCDT Lung test must be considered.

8.1.2 Cost-effectiveness

Our reviews identified two existing cost-effectiveness studies on EarlyCDT Lung, but neither of these studies is considered appropriate due to important differences with the scope of the current decision

problem, including in the patient population, the position and use of EarlyCDT Lung within the diagnostic pathway, and the diagnostic accuracy evidence used to inform it.

We have conducted additional reviews, of diagnostic and screening cost-effectiveness models, to identify value drivers/components of value that could be of relevance to a future assessment of EarlyCDT Lung, and to provide an understanding of the evidence that could be used support such an assessment. The evaluations of diagnostics, like those on EarlyCDT Lung, were supported by little or no empirical evidence on key aspects of model structure and key model parameters. The key mechanism of value used in these studies is consistent with the EarlyCDT Lung studies, and assumes earlier detection (typically in relation to CT surveillance) with stage shift (i.e. identification of cancer at earlier stages of disease) as the key component of value. However, the EAG did not identify evidence supporting the assumption of stage shift for EarlyCDT Lung or underlying any of the diagnostic cost-effectiveness studies.

The review of cost-effectiveness studies of screening strategies showed that the key mechanism of value from the earlier detection of lung cancer is also of stage shift, from the earlier detection of lung cancer. Screening, however, considers the time to diagnosis with clinical identification in relation to the time to diagnosis with the implementation of a screening strategy. The time to diagnosis for solid pulmonary nodules, of concern for the current decision problem, is determined by CT surveillance, and therefore the potential for EarlyCDT Lung to improve time to diagnosis is dictated by the schedule of surveillance scans and the probability of detection at each scan (the earliest of which is at 3 months). The extent of earlier diagnosis expected from a screening strategy (in relation to clinical presentation) is therefore expected to be larger than the extent of earlier diagnosis that could be facilitated by EarlyCDT Lung. However, the evidence on the mean time to pre-clinical progression, often generated within these studies, is currently the best evidence to inform the likelihood of progression under a CT surveillance schedule.

The diagnostic and screening studies, alongside clinical advice, were used together to identify potential value components for EarlyCDT Lung, that could be used to justify the additional costs and health system implications of introducing EarlyCDT Lung in the BTS diagnostic pathway. These include the short-term escalation of diagnostics/treatments and its immediate consequences (such as costs and adverse events and including overtreatment of indolent lesions and the possibility of benign resection) considered alongside the long-term benefits that may arise from earlier identification of lung cancer within the diagnostic pathway.

The reviews highlight that there is currently insufficient evidence to support an explicit quantification of the clinical and economic value of EarlyCDT Lung in the diagnostic pathway of solid pulmonary nodules. A future assessment of EarlyCDT Lung needs to ensure that the evidence supporting the

inclusion and quantification of the abovementioned value components is robust enough to support decision making. Evidence requirements include critical aspects such as the potential for harms of CT surveillance (in terms of delay to diagnosis and the likelihood of stage progression), and the accuracy and clinical utility of EarlyCDT Lung in updating the risk scores commonly used to support management decisions. There is also a lack of epidemiological and service delivery 'intelligence' about pulmonary nodules and their current management in the UK (expected to follow the BTS pathway). This information is essential, not only for supporting future assessments of new technologies in the diagnostic pathway, but also for the prioritisation and planning of further research and development (R&D) efforts and effectiveness/cost-effectiveness research.

We have structured the core components of the decision problem, and conceptualised the implementation of evidence linkage approaches using influence diagrams, which are to be refined as evidence emerges to support a future assessment. These elements were also used to identify further evidence requirements to support an evaluation of the cost-effectiveness of EarlyCDT Lung (or any similar diagnostics) proposed to be used within the BTS pathway.

One of the important aspects emerging regarding conceptualisation of the evidence linkage approaches used to quantify the value of earlier detection is the need for appropriate evidence on the sources and implications of heterogeneity in this patient population. This is particularly relevant as some of these factors may affect the entire evidence linkage pathway, through to outcomes. For example, histology is known to be associated with outcomes, but it is also associated with nodule growth (VDT) which could affect the likelihood of detection over the different time points of CT surveillance, i.e. resulting in different times to (delayed) detection of different histologies. It is therefore important that there is appropriate consideration for these aspects.

8.2 Strengths and limitations of the assessment

This review performed comprehensive searches for EarlyCDT Lung studies. It is likely to have identified all evidence on EarlyCDT Lung currently published, including all studies reported only as conference abstracts. This appears to be the first attempt to synthesise all the evidence on EarlyCDT Lung, including the first meta-analysis for this technology. This review also appears to be the first to attempt to investigate the clinical impact of using EarlyCDT within the BTS diagnostic pathway, although this was limited to a simulation study rather than real data.

The key limitations of the review are a result of the lack of relevant data, the potential for bias in the data that has been published, and its uncertain generalisability to the diagnosis of pulmonary nodules. Consequently, there was little scope for thorough statistical analysis and meta-analysis, and considerably uncertainty as to the robustness of the results.

No direct evidence on the clinical impact of EarlyCDT Lung was identified, severely limiting the ability to investigate how useful or effective EarlyCDT might be in practice. This could only be investigated in a simulation study, which required strong assumptions of uncertain validity. These included strong assumptions on how diagnostic accuracy estimates will translate into post-test risk, and assuming that EarlyCDT Lung is entirely independent of other factors, including nodule size.

8.3 Uncertainties

Uncertainties remain, largely because of the limitations of the data. There appear to be no cohort studies where EarlyCDT Lung is used explicitly within the BTS guidelines (i.e. EarlyCDT Lung being performed after identification of nodules, and in combination with Brock or Herder risk assessment). There is limited data on EarlyCDT Lung in people with pulmonary nodules (only five studies, with only two fully published). All studies are at risk of bias. Consequently, there is too little data in patients with pulmonary nodules to be confident of the diagnostic accuracy of EarlyCDT Lung.

The EAG identified no evidence on the clinical impact of using EarlyCDT Lung, including how patients might be reclassified in terms of risk, or changes of clinical management, and this could only be assessed by simulation. Therefore, the clinical impact of using EarlyCDT within the BTS diagnostic pathway is largely unknown. The EAG identified no relevant evidence on the cost-effectiveness of using EarlyCDT Lung. Therefore, the economic impact of using EarlyCDT within the BTS diagnostic pathway is also largely unknown

The EAG identified comparatively limited evidence on other parts of the BTS diagnostic pathway, including the diagnostic accuracy and clinical impact of Brock and Herder risk assessment, and the clinical impact of CT surveillance. This increases uncertainties as to how using EarlyCDT Lung in the BTS diagnostic pathway might impact patients and health systems.

There is no evidence to support that potential early diagnosis with EarlyCDT Lung will result on stage shift and, importantly, on improved patient outcomes compared to current practice at any of the proposed positionings in the diagnostic pathway. This combined with the limitations of the diagnostic accuracy data on this technology and its limited scope to change clinical decisions on patient management, makes the clinical and economic value of Early CDT Lung highly uncertain.

9 CONCLUSIONS

9.1 Implications for service provision

The EAG concludes that the current evidence on EarlyCDT Lung is insufficient to determine its value in the diagnosis of people with suspect solid pulmonary nodules. This is due to the limited size of the relevant evidence base, uncertainties as to whether current evidence generalises to the UK diagnostic pathway, and a lack of evidence on the diagnostic accuracy and clinical impact of using EarlyCDT Lung.

Based on the limited data available, it appears that EarlyCDT Lung has poor diagnostic accuracy when used in isolation to diagnose pulmonary nodules, with low sensitivity to detect malignancy. It is therefore unclear what it can add to existing diagnostic methods, such as Brock and Herder risk assessment and the use of CT surveillance.

Based on results from the EAG's simulation study, EarlyCDT Lung may have little clinical benefit when diagnosing low-risk or smaller nodules, as it appears unlikely to appropriately change clinical management decisions. EarlyCDT Lung may possibly have clinical value when identifying malignancy in intermediate-risk nodules (10-70% risk after PET-CT scan and Herder risk assessment), by correctly identifying high risk nodules that are malignant, and so might benefit from prompt excision. However, these results are from a simulation study, based on limited data, and requiring various strong assumptions. These conclusions are therefore only suggestions, that would require further research.

The uncertainty over EarlyCDT Lung's clinical utility means its cost-effectiveness is also unclear. Additionally, the main mechanism of value proposed in existing cost-effectiveness studies of EarlyCDT Lung is of earlier detection, but there is no evidence that a meaningful stage shift can happen within the diagnostic pathway for pulmonary nodules. It is, however, clear that a future assessment of EarlyCDT Lung should explore its cost-effectiveness in each of the alternative positionings proposed and in additional subgroups of relevance (such as eligibility for biopsy). This is because the potential for EarlyCDT Lung to alter subsequent clinical management decisions will differ across these groups, and the balance of trade-offs that determine value will also differ (the extent of false positives will determine harm from use of further invasive investigations and the extent of true positives will determine the long term benefits of early detection and treatment). It is also clear that there will need to be appropriate consideration for heterogeneity, as factors such as nodule size or histology determine not only the prevalence of disease, but may also restrict management decisions, be associated with the likelihood of and time to detection, and determine long term outcomes. Any future cost-effectiveness analyses need to appropriately justify the value components included in quantifications, and conduct

extensive sensitivity analyses to explore the impact of the assumptions that are likely to be required given the limitations in the current evidence base,

The EAG notes that these conclusions relate only to the use of EarlyCDT Lung within the BTS recommended pathway for the diagnosis of pulmonary nodules. The EAG has not considered how EarlyCDT Lung might be used in other areas, such as lung cancer screening.

9.2 Suggested research priorities

The EAG's main concern is the general lack of data on EarlyCDT Lung in patients with pulmonary nodules, who will be assessed using BTS guidance. Two studies on EarlyCDT Lung have yet to be fully published. One study, set in China which aims to recruit 1000 patients, is still ongoing,[72] while the other study, the EarlyCDT Lung Cancer Screening study of U.S. patients at high risk of developing lung cancer, has been completed but not yet fully published. It is unclear how many of the patients in these studies have taken the test in a pathway position relevant to where EarlyCDT Lung is most likely to be used in the NHS. It is also unclear whether the test has been used as currently recommended – to update a risk score – or as a simpler positive/negative result.

Large, independent, prospective cohort studies are therefore needed, where EarlyCDT Lung is used in patients with identified pulmonary nodules. Patients should be diagnosed and manged in line with the BTS diagnostic pathway (Figure 1), with sufficient follow up to confirm malignancy by biopsy or surgery, or its absence with at least 2 years' follow-up without nodule growth. This will permit the estimation of the diagnostic accuracy of EarlyCDT Lung:

- In isolation
- In combination with Brock risk
- In combination with PET-CT and Herder risk (in patients receiving a PET-CT scan)

These cohort studies should also assess the clinical impact of EarlyCDT Lung by reporting outcomes including:

- Impact on risk classification (e.g. moving from low risk (<10%) to intermediate risk (10-70%), or intermediate risk to high risk (>70%))
- Change in clinical management (e.g. moving from CT surveillance to biopsy, or biopsy to immediate excision)
- Timing and tumour stage at detection and treatment of malignant nodules
- Avoidance of unnecessary CT or PET-CT scans
- Promotion of unnecessary PET-CT scans, biopsies or surgical excisions (as a consequence of false-positive EarlyCDT Lung) and their consequent risk of adverse events

The EAG has concerns that the proposed risk model for EarlyCDT Lung (Figure 2) may be based on biased estimates of diagnostic accuracy (see Section 3.1.7.3). This risk model requires proper validation in independent cohorts. The cohort studies described above could be used for this. If the model is found not to be valid (i.e. its estimated risks do not match observed risks) a new model will be required, based on robust diagnostic accuracy data from new cohorts. Further cohort studies would then be required to validate the new model

Diagnostic accuracy studies do not tell us whether differences in accuracy result in clinically important effects on patient health outcomes. These effects may occur as a result of changes to further therapeutic or diagnostic interventions, based on test results. The optimal approach to determining the clinical value of EarlyCDT Lung would therefore be to conduct a randomised controlled trial, where patients with identified pulmonary nodules are randomised either to standard BTS management (Figure 1) or to BTS management with EarlyCDT Lung included (as in Figure 3). A trial may be beneficial if cohort studies suggest potential, but inconclusive, benefits of EarlyCDT Lung. A randomised trial may not be required if evidence from high-quality cohort studies is sufficient to support the use of EarlyCDT Lung.

Currently, the broader evidence base on the whole BTS diagnostic pathway is insufficient to allow explicit and formal quantifications of the clinical end economic value of EarlyCDT Lung (or any other future test in this area). Although the EAG has not conducted a full systematic review of the entire pathway, our limited review, and the reviews conducted to support the BTS guidance, both suggest that large well-designed and UK-based prospective cohort studies are particularly needed to investigate the following:

- The diagnostic accuracy and clinical impact of using the Brock risk model, in the context of UK clinical practice;
- The diagnostic accuracy and clinical impact of using PET-CT scans and the Herder risk model, in the context of UK clinical practice;
- The clinical consequences of CT surveillance (e.g. number of cancers identified and missed, delay in diagnosis and the possibility of tumour progression); and
- How patient and nodule characteristics determine malignancy prevalence, eligibility for alternative clinical management options, likelihood and time to detection under CT surveillance, and patient outcomes.

To further support the evidence linkage approaches likely to be required to support a future cost-effectiveness study (in the absence of an outcomes study), a number of additional studies could be important. These could include a comparative analysis of the pre-clinical progression models developed from screening studies which would provide a broader understanding of the speed of pre-clinical progression of lung cancer. Additionally, evidence allowing a better understanding of current

variation in management decisions in the intermediate risk group and of its determinants (malignancy risk, patient preference and fitness to undergo invasive procedures), would also be valuable to allow appropriately reflecting this variation in a future decision model. Finally, evidence on the current extent of benign resection be important.

10 REFERENCES

1. NHS. *NHS Lung cancer overview*. 2019. URL: https://www.nhs.uk/conditions/lung-cancer/ (accessed 16/08/21.
2. National Institute for Health and Care Excellence. *Lung cancer: diagnosis and management. NICE guideline [NG122]*. NICE; 2019. URL: https://www.nice.org.uk/guidance/ng122 (accessed 2nd February 2021).
3. Callister MEJ, Baldwin DR, Akram AR, Barnard S, Cane P, Draffan J*, et al.* British Thoracic Society guidelines for the investigation and management of pulmonary nodules: accredited by NICE. *Thorax* 2015;**70**:ii1-54. 10.1136/thoraxjnl-2015-207168
4. Bueno J, Landeras L, Chung JH. Updated Fleischner Society Guidelines for managing incidental pulmonary nodules: common questions and challenging scenarios. *Radiographics* 2018;**38**:1337-50. 10.1148/rg.2018180017
5. Gould MK, Donington J, Lynch WR, Mazzone PJ, Midthun DE, Naidich DP*, et al.* Evaluation of individuals with pulmonary nodules: when is it lung cancer? Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines.
6. Al-Ameri A, Malhotra P, Thygesen H, Plant PK, Vaidyanathan S, Karthik S*, et al.* Risk of malignancy in pulmonary nodules: a validation study of four prediction models. *Lung Cancer* 2015;**89**:27-30. 10.1016/j.lungcan.2015.03.018
7. Chapman CJ, Healey G, Murray A, Peek LJ, Hamilton-Fairley G, Robertson JFR. Improved specificity of the EarlyCDT-Lung test through additional autoantibody assays. Paper presented at: American Thoracic Society 2012 International Conference; Denver, May 18-23.
8. Jett JR, Dyer D, Kern J, Rollins D, Phillips M. Screening for lung cancer with the EarlyCDT-Lung and computed tomography. *J Thorac Oncol* 2015;**10**:S306.
9. Isbell JM, Deppen S, Putnam JB, Nesbitt JC, Lambright ES, Dawes A*, et al.* Existing General Population Models Inaccurately Predict Lung Cancer Risk in Patients Referred for Surgical Evaluation. *Ann Thorac Surg* 2011;**91**:227-33. 10.1016/j.athoracsur.2010.08.054
10. Healey GF, Macdonald IK, Reynolds C, Allen J, Murray A. Tumor-associated autoantibodies: re-optimization of EarlyCDT-Lung diagnostic performance and its application to indeterminate pulmonary nodules. *Journal of Cancer Therapy* 2017;**8**:506-17. 10.4236/jct.2017.85043
11. Herder GJ, van Tinteren H, Golding RP, Kostense PJ, Comans EF, Smit EF*, et al.* Clinical prediction model to characterize pulmonary nodules: validation and added value of 18 F-fluorodeoxyglucose positron emission tomography. *Chest* 2005;**128**:2490-6. https://doi.org/10.1378/chest.128.4.2490
12. Healey G, Murray A, Peek L, Fredericks L. Measuring the improvement in lung nodule cancer risk assessment when EarlyCDT-lung is applied to a clinically-tested high-risk cohort. *Am J Respir Crit Care Med* 2015;**191**:A3585.
13. Sullivan FM, Mair FS, Anderson W, Armory P, Briggs A, Chew C*, et al.* Earlier diagnosis of lung cancer in a randomised trial of an autoantibody blood test followed by imaging. *Eur Respir J* 2021;**57**:2000670. https://dx.doi.org/10.1183/13993003.00670-2020
14. Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care*. York: CRD, University of York; 2009.
15. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic reviews and Meta-Analyses: the PRISMA Statement. *PLoS Med* 2009;**6**:e1000097. 10.1371/journal.pmed.1000097
16. Yang Y, Abel L, Buchanan J, Fanshawe T, Shinkins B. Use of decision modelling in economic evaluations of diagnostic tests: an appraisal and review of Health Technology Assessments in the UK. *Pharmacoecon Open* 2019;**3**:281-91. 10.1007/s41669-018-0109-9
17. Canadian Agency for Drugs and Technologies in Health. *CADTH's database search filters - narrow economic filter*. URL: https://www.cadth.ca/strings-attached-cadths-database-search-filters (accessed 22nd March 2021).
18. Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M. Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–2. *Med Decis Making* 2012;**32**:678-89.

19. Kaltehthaler E, Tappenden P, Paisley S, Squires H. *NICE DSU Technical support document 13: Identifying and reviewing evidence to inform the conceptualisation and population of cost-effectiveness models*. Sheffield: Decision Support Unit, ScHARR, University of Sheffield; 2011.
20. National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal 2013*: NICE; 2013.
21. Kjaerulff UB, Madsen AL. *Bayesian networks and influence diagrams: a guide to construction and analysis*. New York: Springer; 2008.
22. Owens DK, Shachter RD, Nease RF. Representation and analysis of medical decision problems with influeance diagrams. *Med Decis Making* 1997;**17**:241-62. 10.1177/0272989x9701700301
23. Soares MO, Walker S, Palmer SJ, Sculpher MJ. Establishing the value of diagnostic and prognostic tests in Health Technology Assessment. *Med Decis Making* 2018;**38**:495-508.
24. Boyle P, Chapman CJ, Holdenrieder S, Murray A, Robertson C, Wood WC*, et al.* Clinical validation of an autoantibody test for lung cancer. *Ann Oncol* 2011;**22**:383-9. https://dx.doi.org/10.1093/annonc/mdq361
25. Chapman CJ, Murray A, McElveen JE, Sahin U, Luxemburger U, Tureci O*, et al.* Autoantibodies in lung cancer: possibilities for early detection and subsequent cure. *Thorax* 2008;**63**:228-33.
26. Chapman CJ, Thorpe AJ, Murray A, Parsy-Kowalska CB, Allen J, Stafford KM*, et al.* Immunobiomarkers in small cell lung cancer: potential early cancer signals. *Clin Cancer Res* 2011;**17**:1474-80. https://dx.doi.org/10.1158/1078-0432.CCR-10-1363
27. Lam S, Boyle P, Healey GF, Maddison P, Peek L, Murray A*, et al.* EarlyCDT-Lung: an immunobiomarker test as an aid to early detection of lung cancer. *Cancer Prev Res (Phila)* 2011;**4**:1126-34. https://dx.doi.org/10.1158/1940-6207.CAPR-10-0328
28. Macdonald IK, Allen J, Murray A, Parsy-Kowalska CB, Healey GF, Chapman CJ*, et al.* Development and validation of a high throughput system for discovery of antigens for autoantibody detection. *PLoS ONE* 2012;**7**:e40759. https://dx.doi.org/10.1371/journal.pone.0040759
29. Macdonald IK, Murray A, Healey GF, Parsy-Kowalska CB, Allen J, McElveen J*, et al.* Application of a high throughput method of biomarker discovery to improvement of the EarlyCDT-Lung test. *PLoS ONE* 2012;**7**:e51002. https://dx.doi.org/10.1371/journal.pone.0051002
30. Murray A, Chapman CJ, Healey G, Peek LJ, Parsons G, Baldwin D*, et al.* Technical validation of an autoantibody test for lung cancer. *Ann Oncol* 2010;**21**:1687-93. https://dx.doi.org/10.1093/annonc/mdp606
31. Chapman CJ, Healey GF, Murray A, Boyle P, Robertson C, Peek LJ*, et al.* EarlyCDT-Lung test: improved clinical utility through additional autoantibody assays. *Tumour Biol* 2012;**33**:1319-26. https://dx.doi.org/10.1007/s13277-012-0379-2
32. Healey GF, Lam S, Boyle P, Hamilton-Fairley G, Peek LJ, Robertson JF. Signal stratification of autoantibody levels in serum samples and its application to the early detection of lung cancer. *J Thorac Dis* 2013;**5**:618-25. https://dx.doi.org/10.3978/j.issn.2072-1439.2013.08.65
33. Jett JR, Peek LJ, Fredericks L, Jewell W, Pingleton WW, Robertson JF. Audit of the autoantibody test, EarlyCDT-Lung, in 1600 patients: an evaluation of its performance in routine clinical practice. *Lung Cancer* 2014;**83**:51-5. https://dx.doi.org/10.1016/j.lungcan.2013.10.008
34. Massion PP, Healey GF, Peek LJ, Fredericks L, Sewell HF, Murray A*, et al.* Autoantibody signature enhances the positive predictive power of computed tomography and nodule-based risk models for detection of lung cancer. *J Thorac Oncol* 2017;**12**:578-84. https://dx.doi.org/10.1016/j.jtho.2016.08.143
35. Chapman CJ, Murray A, Healey G, Hamilton F, Robertson JFR. Early identification of lung cancer using autoantibody assays. *Tumour Biol* 2012;**33**:S41. http://dx.doi.org/10.1007/s13277-012-0515-z
36. Healey G, Peek L, Fredericks L, Hamilton-Fairley G, Robertson J. The additive effect of the stratified EarlyCDT-Lung serum autoantibody test on the assessment of lung cancer risk. *Chest* 2012;**142**:638A. http://dx.doi.org/10.1378/chest.1390402
37. Healey GF, Chapman CJ, Murray A, Hamilton-Fairley GN, Robertson JF. Signal stratification of autoantibody levels in serum samples and its application to the early diagnosis of lung cancer. *Am J Respir Crit Care Med* 2012;**185**:A6365.

38. Jett J, Healey G, Murray A, Hamilton-Fairley G, Fraser S, Peek L*, et al.* Potential utility of a positive EarlyCDT-Lung blood biomarker test in indeterminate pulmonary nodules. *J Thorac Oncol* 2018;**13**:S782. http://dx.doi.org/10.1016/j.jtho.2018.08.1356
39. Jett JR, Jewell W, Robertson JF. EarlyCDT-Lung test: audit of the first 1000 patients in clinical practice. *J Thorac Oncol* 2011;**6**:S1387-S8.
40. Jett JR, Peek LJ, Fredericks L, Jewell W, Pingleton WW, Robertson JFR. An audit of the clinical performance of EarlyCDT-Lung. *J Thorac Oncol* 2012;**7**:S222. http://dx.doi.org/10.1097/JTO.0b013e318269fc07
41. Kucera R. Benefits of using EarlyCDT-Lung in clinical practice. *Chest* 2012;**142**:637A. http://dx.doi.org/10.1378/chest.1390390
42. Massion P, Peek LJ, Fredericks L, Jewell W, Pingleton WW, Robertson JFR. Autoantibodies to a panel of lung cancer-associated antigens can provide significant discrimination between malignant and non-malignant lung nodules. *J Thorac Oncol* 2013;**8**:S979. http://dx.doi.org/10.1097/01.JTO.0000438438.14562.c8
43. Peek L, Fredericks L, Jewell W, Pingleton W, Robertson J. EarlyCDT-lung enhances the positive predictive power of computed tomography for detecting lung cancer. *Chest* 2012;**142**:560A. http://dx.doi.org/10.1378/chest.1390243
44. Spiegler P. Serum autoantibody testing for lung cancer. *Clin Pulm Med* 2014;**21**:241-2. http://dx.doi.org/10.1097/CPM.0000000000000050
45. National Jewish Health. *Lung cancer screening study with low-dose CT scan and blood biomarker*. ClinicalTrials.gov. Bethesda (MD): National Library of Medicine (US); 2012. URL: https://ClinicalTrials.gov/show/NCT01700257 (accessed 9th March 2021).
46. Jett J, Dyer D, Kern J, Rollins D, Phillips M, Finigan J. Screening for lung cancer with EarlyCDT-Lung blood biomarkers and computed tomography. *J Thorac Oncol* 2017;**12**:S569-70.
47. Ohillips M, Rollins D, Dyer D, Kern J, Jett J, Finigan J. Analysis of the early CDT-blood biomarker for lung cancer in higher vs. lower risk cohorts. *J Thorac Oncol* 2017;**12**:S591.
48. Clark ME, Bedford LE, Young B, Robertson JFR, das Nair R, Vedhara K*, et al.* Lung cancer CT screening: psychological responses in the presence and absence of pulmonary nodules. *Lung Cancer* 2018;**124**:160-7. https://dx.doi.org/10.1016/j.lungcan.2018.08.001
49. Clark ME, Young B, Bedford LE, das Nair R, Robertson JFR, Vedhara K*, et al.* Lung cancer screening: does pulmonary nodule detection affect a range of smoking behaviours? *J Public Health (Oxf)* 2019;**41**:600-8. https://dx.doi.org/10.1093/pubmed/fdy158
50. Sullivan FM, Farmer E, Mair FS, Treweek S, Kendrick D, Jackson C*, et al.* Detection in blood of autoantibodies to tumour antigens as a case-finding method in lung cancer using the EarlyCDT-Lung Test (ECLS): study protocol for a randomized controlled trial. *BMC Cancer* 2017;**17**:187. https://dx.doi.org/10.1186/s12885-017-3175-y
51. Young B, Vedhara K, Kendrick D, Littleford R, Robertson JFR, Sullivan FM*, et al.* Determinants of motivation to quit in smokers screened for the early detection of lung cancer: a qualitative study. *BMC Public Health* 2018;**18**:1276. https://dx.doi.org/10.1186/s12889-018-6211-1
52. Bedford LE. *The emotional impact of screening for lung cancer.* PhD. Nottingham: University of Nottingham; 2017.
53. Clark M, Bedford L, Young B, Das Nair R, Robertson J, Vedhara K*, et al.* Is lung cancer screening associated with a negative psychological impact? *J Thorac Oncol* 2017;**12**:S1760.
54. Dorward A, Frances Mair F, Sullivan F, Vedhara K, Kendrick D, Treweek S*, et al.* A randomised controlled study of lung cancer screening in Scotland using the detection of autoantibodies to tumour antigens (EarlyCDT-Lung Test). *Thorax* 2016;**71**:A78-A9. http://dx.doi.org/10.1136/thoraxjnl-2016-209333.138
55. Sullivan F, Dorward A, Mair F, Schembri S, Robertson J, Vedhara K*, et al.* A new partnership with biomarker research: a randomised controlled trial of the detection in blood of autoantibodies to tumour antigens as a case-finding method in lung cancer using the EarlyCDT-Lung test in Scotland (ECLS). *Eur J Cancer Care* 2014;**23**:31.
56. Sullivan F, Schembri S. Progress with an RCT of the detection of autoantibodies to tumour antigens in lung cancer using the early CDT-Lung test in Scotland (ECLS). *J Thorac Oncol* 2015;**10**:S306.

57. Sullivan F, Dorward A, Mair F, Schembri S, Vedhara K, Kendrick D*, et al.* An RCT of the detection of autoantibodies to tumor antigens in lung cancer using the EarlyCDT-Lung test in Scotland (ECLS) in 12 208 study subjects. *J Thorac Oncol* 2017;**12**:S1095.
58. Sullivan F, Schembri S. An interim report on the detection in blood of autoantibodies to tumour antigens as a case-finding method in lung cancer using the EarlyCDT-Lung Test in Scotland (ECLS). *Eur Respir J* 2017;**50**:OA1471. http://dx.doi.org/10.1183/1393003.congress-2017.OA1471
59. Sullivan F, Schembri S. Early detection of cancer of the lung scotland (ECLS): trial results. *J Thorac Oncol* 2019;**14**:S5. http://dx.doi.org/10.1016/j.jtho.2019.08.056
60. Young B, Bedford L, Vedhara K, das Nair R, Robertson J, Kendrick D. Changes in smoking behavior in the early cancer detection test lung cancer Scotland (ECLS) study. *J Thorac Oncol* 2017;**12**:S1878.
61. Young B, Vedhara K, Kendrick D, Robertson J, Das N. How does screening for the early detection of lung cancer facilitate smoking cessation? A qualitative study of screened smokers. *J Thorac Oncol* 2017;**12**:S2297.
62. Young B, Vedhara K, Kendrick D, Robertson J, das N. Decisions about smoking in patients screened with the EarlyCDT-Lung test for the early detection of lung cancer: a qualitative study. *J Epidemiol Community Health* 2017;**71**:A19. 10.1136/jech-2017-SSMAbstracts.36
63. Sullivan F, Schembri S. Erratum to: Sullivan F, Schembri S. Early detection of cancer of the lung Scotland (ECLS): trial results [abstract]. J Thorac Oncol. 2019;14:10(S):S5. *J Thorac Oncol* 2020;**15**:146.
64. Gonzalez Maldonado S, Johnson T, Motsch E, Delorme S, Kaaks R. Can autoantibody tests enhance lung cancer screening?-an evaluation of EarlyCDT-Lung in context of the German Lung Cancer Screening Intervention Trial (LUSI). *Transl Lung Cancer Res* 2021;**10**:233-42. https://dx.doi.org/10.21037/tlcr-20-727
65. Lin G, Tripathi P, Naemi H, Nair V. The performance of two commercial, diagnostic blood tests to evaluate the incidentally detected, indeterminate pulmonary nodule. *Chest* 2016;**150**:733A. http://dx.doi.org/10.1016/j.chest.2016.08.828
66. Lau CL. A pilot study of the autoantibodies to tumor antigens in lung cancer using the EarlyCDT-lung test in Hong Kong. *J Thorac Oncol* 2017;**12**:S2323.
67. Chu GCW, Lazare K, Sullivan F. Serum and blood based biomarkers for lung cancer screening: a systematic review. *BMC Cancer* 2018;**18**:181. https://dx.doi.org/10.1186/s12885-018-4024-3
68. Qin J, Zeng N, Yang T, Wan C, Chen L, Shen Y*, et al.* Diagnostic value of autoantibodies in lung cancer: a systematic review and meta-analysis. *Cell Physiol Biochem* 2018;**51**:2631-46. https://dx.doi.org/10.1159/000495935
69. Tang ZM, Ling ZG, Wang CM, Wu YB, Kong JL. Serum tumor-associated autoantibodies as diagnostic biomarkers for lung cancer: a systematic review and meta-analysis. *PLoS ONE* 2017;**12**:e0182117. https://dx.doi.org/10.1371/journal.pone.0182117
70. Yang B, Li X, Ren T, Yin Y. Autoantibodies as diagnostic biomarkers for lung cancer: a systematic review. *Cell Death Discovery* 2019;**5**:126. https://dx.doi.org/10.1038/s41420-019-0207-1
71. Kelly KE, Culbertson WH. Stage 1a non-small cell lung cancer detected by assaying autoantibodies to tumor antigens. *Chest* 2010;**138**:246A. http://dx.doi.org/10.1378/chest.10544
72. Chunxue B. *Early cancer detection test - lung cancer China*. ClinicalTrials.gov. Bethesda (MD): National Library of Medicine (US); 2020. URL: https://ClinicalTrials.gov/show/NCT04216511 (accessed 9th March 2021).
73. Borg M, Wen SWC, Nederby L, Hansen TF, Jakobsen A, Andersen RF*, et al.* Performance of the EarlyCDT® Lung test in detection of lung cancer and pulmonary metastases in a high-risk cohort. *Lung Cancer* 2021;**158**:85-90. https://doi.org/10.1016/j.lungcan.2021.06.010
74. Al-Ameri A, Malhotra P, Thygesen H, Vaidyanathan S, Karthik S, Scarsbrook A*, et al.* S74 Assessing the diagnostic accuracy of the British Thoracic Society algorithm for investigation of solid pulmonary nodules. *Thorax* 2015;**70**:A44. 10.1136/thoraxjnl-2015-207770.80
75. Guo H-Y, Lin J-T, Huang H-H, Gao Y, Yan M-R, Sun M*, et al.* Development and Validation of a 18F-FDG PET/CT-Based Clinical Prediction Model for Estimating Malignancy in Solid Pulmonary Nodules Based on a Population With High Prevalence of Malignancy. *Clinical Lung Cancer* 2020;**21**:47-55. https://doi.org/10.1016/j.cllc.2019.07.014

76. Winter A, Aberle DR, Hsu W. External validation and recalibration of the Brock model to predict probability of cancer in pulmonary nodules using NLST data. *Thorax* 2019;**74**:551-63. https://dx.doi.org/10.1136/thoraxjnl-2018-212413
77. Perandini S, Soardi GA, Larici AR, Del Ciello A, Rizzardi G, Solazzo A*, et al.* Multicenter external validation of two malignancy risk prediction models in patients undergoing 18F-FDG-PET for solitary pulmonary nodule evaluation. *Eur Radiol* 2017;**27**:2042-6. https://dx.doi.org/10.1007/s00330-016-4580-3
78. Basso Dias A, Zanon M, Altmayer S, Sartori Pacini G, Henz Concatto N, Watte G*, et al.* Fluorine 18-FDG PET/CT and diffusion-weighted MRI for malignant versus benign pulmonary lesions: a meta-analysis. *Radiology* 2019;**290**:525-34. https://dx.doi.org/10.1148/radiol.2018181159
79. Divisi D, Barone M, Bertolaccini L, Zaccagna G, Gabriele F, Crisci R. Diagnostic performance of fluorine-18 fluorodeoxyglucose positron emission tomography in the management of solitary pulmonary nodule: a meta-analysis. *J Thorac Dis* 2018;**10**:S779-S89.
80. Jia Y, Gong W, Zhang Z, Tu G, Li J, Xiong F*, et al.* Comparing the diagnostic value of 18F-FDG-PET/CT versus CT for differentiating benign and malignant solitary pulmonary nodules: a meta-analysis. *J Thorac Dis* 2019;**11**:2082-98.
81. Li ZZ, Huang YL, Song HJ, Wang YJ, Huang Y. The value of 18F-FDG-PET/CT in the diagnosis of solitary pulmonary nodules: a meta-analysis. *Medicine* 2018;**97**:e0130. https://dx.doi.org/10.1097/MD.0000000000010130
82. Madsen PH, Holdgaard PC, Christensen JB, Hoilund-Carlsen PF. Clinical utility of F-18 FDG PET-CT in the initial evaluation of lung cancer. *Eur J Nucl Med Mol Imaging* 2016;**43**:2084-97. https://dx.doi.org/10.1007/s00259-016-3407-4
83. Ruilong Z, Daohai X, Li G, Xiaohong W, Chunjie W, Lei T. Diagnostic value of 18F-FDG-PET/CT for the evaluation of solitary pulmonary nodules: a systematic review and meta-analysis. *Nucl Med Commun* 2017;**38**:67-75.
84. Wang HQ, Zhao L, Zhao J, Wang Q. Analysis on early detection of lung cancer by PET/CT scan. *Asian Pac J Cancer Prev* 2015;**16**:2215-7.
85. Zhang L, Wang Y, Lei J, Tian J, Zhai Y. Dual time point 18FDG-PET/CT versus single time point 18FDG-PET/CT for the differential diagnosis of pulmonary nodules: a meta-analysis. *Acta Radiol* 2013;**54**:770-7. https://dx.doi.org/10.1177/0284185113481594
86. Zhao H, Marshall HM, Yang IA, Bowman RV, Ayres J, Crossin J*, et al.* Screen-detected subsolid pulmonary nodules: long-term follow-up and application of the PanCan lung cancer risk prediction model. *Br J Radiol* 2016;**89**:20160016. https://dx.doi.org/10.1259/bjr.20160016
87. Lococo F, Muoio B, Chiappetta M, Nachira D, Petracca Ciavarella L, Margaritora S*, et al.* Diagnostic performance of PET or PET/CT with different radiotracers in patients with suspicious lung cancer or pleural tumours according to published meta-analyses. *Contrast Media Mol Imaging* 2020;**2020**:5282698.
88. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB*, et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;**155**:529-36.
89. Evangelista L, Cuocolo A, Pace L, Mansi L, Del Vecchio S, Miletto P*, et al.* Performance of FDG-PET/CT in solitary pulmonary nodule based on pre-test likelihood of malignancy: results from the ITALIAN retrospective multicenter trial. *Eur J Nucl Med Mol Imaging* 2018;**45**:1898-907. https://dx.doi.org/10.1007/s00259-018-4016-1
90. Spadafora M, Evangelista L, Fiordoro S, Porcaro F, Sicignano M, Mansi L. The multicenter Italian trial assesses the performance of FDG-PET /CT related to pre-test cancer risk in patients with solitary pulmonary nodules and introduces a segmental thoracic diagnostic strategy. *Curr Radiopharm* 2020;**13**:243-8. https://dx.doi.org/10.2174/1874471013666200318142210
91. Weir-McCall JR, Harris S, Miles KA, Qureshi NR, Rintoul RC, Dizdarevic S*, et al.* Impact of solitary pulmonary nodule size on qualitative and quantitative assessment using 18F-fluorodeoxyglucose PET/CT: the SPUTNIK trial. *Eur J Nucl Med Mol Imaging* 2021;**48**:1560-9.
92. Crosbie PA, Balata H, Evison M, Atack M, Bayliss-Brideaux V, Colligan D*, et al.* Implementing lung cancer screening: baseline results from a community-based 'Lung Health Check' pilot in deprived areas of Manchester. *Thorax* 2019;**74**:405-9.

93. Crosbie PA, Balata H, Evison M, Atack M, Bayliss-Brideaux V, Colligan D*, et al.* Second round results from the Manchester 'Lung Health Check' community-based targeted lung cancer screening pilot. *Thorax* 2019;**74**:700-4.
94. Field JK, Duffy SW, Baldwin DR, Whynes DK, Devaraj A, Brain KE*, et al.* UK Lung Cancer RCT Pilot Screening Trial: baseline findings from the screening arm provide evidence for the potential implementation of lung cancer screening. *Thorax* 2016;**71**:161-70.
95. Walter JE, Heuvelmans MA, Ten Haaf K, Vliegenthart R, van der Aalst CM, Yousaf-Khan U*, et al.* Persisting new nodules in incidence rounds of the NELSON CT lung cancer screening study. *Thorax* 2019;**74**:247-53.
96. Bartlett EC, Kemp SV, Ridge CA, Desai SR, Mirsadraee S, Morjaria JB*, et al.* Baseline results of the West London lung cancer screening pilot study - Impact of mobile scanners and dual risk model utilisation. *Lung Cancer* 2020;**148**:12-9.
97. Becker N, Motsch E, Trotter A, Heussel CP, Dienemann H, Schnabel PA*, et al.* Lung cancer mortality reduction by LDCT screening - results from the randomized German LUSI trial. *Int J Cancer* 2020;**146**:1503-13.
98. Liu GS, Wang SQ, Liu HL, Liu Y, Fu YF, Shi YB. Computed tomography-guided biopsy for small (<=20 mm) lung nodules: a meta-analysis. *J Comput Assist Tomogr* 2020;**44**:841-6. https://dx.doi.org/10.1097/RCT.0000000000001071
99. Yan GW, Bhetuwal A, Yan GW, Sun QQ, Niu XK, Zhou Y*, et al.* A systematic review and meta-analysis of C-arm cone-beam CT-guided percutaneous transthoracic needle biopsy of lung nodules. *Pol J Radiol* 2017;**82**:152-60. https://dx.doi.org/10.12659/PJR.899626
100. Yang JS, Liu YM, Mao YM, Yuan JH, Yu WQ, Cheng RD*, et al.* Meta-analysis of CT-guided transthoracic needle biopsy for the evaluation of the ground-glass opacity pulmonary lesions. *Br J Radiol* 2014;**87**:20140276. https://dx.doi.org/10.1259/bjr.20140276
101. Zhan P, Zhu QQ, Miu YY, Liu YF, Wang XX, Zhou ZJ*, et al.* Comparison between endobronchial ultrasound-guided transbronchial biopsy and CT-guided transthoracic lung biopsy for the diagnosis of peripheral lung cancer: a systematic review and meta-analysis. *Transl Lung Cancer Res* 2017;**6**:23-34. https://dx.doi.org/10.21037/tlcr.2017.01.01
102. Zhang HF, Zeng XT, Xing F, Fan N, Liao MY. The diagnostic accuracy of CT-guided percutaneous core needle biopsy and fine needle aspiration in pulmonary lesions: a meta-analysis. *Clin Radiol* 2016;**71**:e1-10. https://dx.doi.org/10.1016/j.crad.2015.09.009
103. Han Y, Kim HJ, Kong KA, Kim SJ, Lee SH, Ryu YJ*, et al.* Diagnosis of small pulmonary lesions by transbronchial lung biopsy with radial endobronchial ultrasound and virtual bronchoscopic navigation versus CT-guided transthoracic needle biopsy: a systematic review and meta-analysis. *PLoS One* 2018;**13**:e0191590. https://dx.doi.org/10.1371/journal.pone.0191590
104. Sryma PB, Mittal S, Madan NK, Tiwari P, Hadda V, Mohan A*, et al.* Efficacy of Radial Endobronchial Ultrasound (R-EBUS) guided transbronchial cryobiopsy for peripheral pulmonary lesions (PPL's): a systematic review and meta-analysis. *Pulmonology* 2021. https://doi.org/10.1016/j.pulmoe.2020.12.006
105. Ali MS, Sethi J, Taneja A, Musani A, Maldonado F. Computed tomography bronchus sign and the diagnostic yield of guided bronchoscopy for peripheral pulmonary lesions. A systematic review and meta-analysis. *Ann Am Thorac Soc* 2018;**15**:978-87. https://dx.doi.org/10.1513/AnnalsATS.201711-856OC
106. Edelsberg J, Weycker D, Atwood M, Hamilton-Fairley G, Jett JR. Cost-effectiveness of an autoantibody test (EarlyCDT-Lung) as an aid to early diagnosis of lung cancer in patients with incidentally detected pulmonary nodules. *PLoS ONE* 2018;**13**:e0197826. https://dx.doi.org/10.1371/journal.pone.0197826
107. 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**:e0237492. https://dx.doi.org/10.1371/journal.pone.0237492
108. Tanner NT, Porter A, Gould MK, Li XJ, Vachani A, Silvestri GA. Physician assessment of pretest probability of malignancy and adherence with guidelines for pulmonary nodule evaluation. *Chest* 2017;**152**:263-70. 10.1016/j.chest.2017.01.018

109. Gould MK, Sanders GD, Barnett PG, Rydzak CE, Maclean CC, McClellan MB, *et al.* Cost-effectiveness of alternative management strategies for patients with solitary pulmonary nodules. *Ann Intern Med* 2003;**138**:724-35.
110. Steele JD, Buell P. Asymptomatic solitary pulmonary nodules. Host survival, tumor size, and growth rate. *J Thorac Cardiovasc Surg* 1973;**65**:140-51.
111. Black WC, Gareen IF, Soneji SS, Sicks JD, Keeler EB, Aberle DR, *et al.* Cost-effectiveness of CT screening in the National Lung Screening Trial. *N Engl J Med* 2014;**371**:1793-802. 10.1056/NEJMoa1312547
112. D'Andrea E, Choudhry NK, Raby B, Weinhouse GL, Najafzadeh M. A bronchial-airway gene-expression classifier to improve the diagnosis of lung cancer: clinical outcomes and cost-effectiveness analysis. *Int J Cancer* 2020;**146**:781-90. https://dx.doi.org/10.1002/ijc.32333
113. Deppen SA, Davis WT, Green EA, Rickman O, Aldrich MC, Fletcher S, *et al.* Cost-effectiveness of initial diagnostic strategies for pulmonary nodules presenting to thoracic surgeons. *Ann Thorac Surg* 2014;**98**:1214-22. https://dx.doi.org/10.1016/j.athoracsur.2014.05.025
114. Dietlein M, Weber K, Gandjour A, Moka D, Theissen P, Lauterbach KW, *et al.* Cost-effectiveness of FDG-PET for the management of solitary pulmonary nodules: a decision analysis based on cost reimbursement in Germany. *Eur J Nucl Med* 2000;**27**:1441-56.
115. Goehler A, McMahon PM, Lumish HS, Wu CC, Munshi V, Gilmore M, *et al.* Cost-effectiveness of follow-up of pulmonary nodules incidentally detected on cardiac computed tomographic angiography in patients with suspected coronary artery disease. *Circulation* 2014;**130**:668-75. https://dx.doi.org/10.1161/CIRCULATIONAHA.113.007306
116. Jiang B, Linden PA, Gupta A, Jarrett C, Worrell SG, Ho VP, *et al.* Conventional Computed Tomographic Calcium Scoring vs full chest CTCS for lung cancer screening: a cost-effectiveness analysis. *BMC Pulm Med* 2020;**20**:187. https://dx.doi.org/10.1186/s12890-020-01221-8
117. Lejeune C, Al Zahouri K, Woronoff-Lemsi MC, Arveux P, Bernard A, Binquet C, *et al.* Use of a decision analysis model to assess the medicoeconomic implications of FDG PET imaging in diagnosing a solitary pulmonary nodule. *Eur J Health Econ* 2005;**50**:203-14.
118. Rickets W, Lau KKW, Pollit V, Mealing S, Leonard C, Mallender P, *et al.* Exploratory cost-effectiveness model of electromagnetic navigation bronchoscopy (ENB) compared with CT-guided biopsy (TTNA) for diagnosis of malignant indeterminate peripheral pulmonary nodules. *BMJ Open Respiratory Research* 2020;**7**:e000595. https://dx.doi.org/10.1136/bmjresp-2020-000595
119. Allen BD, Schiebler ML, Sommer G, Kauczor HU, Biederer J, Kruser TJ, *et al.* Cost-effectiveness of lung MRI in lung cancer screening. *Eur Radiol* 2020;**30**:1738-46. https://dx.doi.org/10.1007/s00330-019-06453-9
120. Caro JJ, Klittich WS, Strauss G. Could chest X-ray screening for lung cancer be cost-effective? *Cancer* 2000;**89**:2502-5.
121. Chirikos TN, Hazelton T, Tockman M, Clark R. Screening for lung cancer with CT: a preliminary cost-effectiveness analysis. *Chest* 2002;**121**:1507-14.
122. Cressman S, Peacock SJ, Tammemagi MC, Evans WK, Leighl NB, Goffin JR, *et al.* The cost-effectiveness of high-risk lung cancer screening and drivers of program efficiency. *J Thorac Oncol* 2017;**12**:1210-22. https://dx.doi.org/10.1016/j.jtho.2017.04.021
123. Du Y, Sidorenkov G, Heuvelmans MA, Groen HJM, Vermeulen KM, Greuter MJW, *et al.* Supplementary data for a model-based health economic evaluation on lung cancer screening with low-dose computed tomography in a high-risk population. *Data Brief* 2020;**31**:105999. https://dx.doi.org/10.1016/j.dib.2020.105999
124. Du Y, Sidorenkov G, Heuvelmans MA, Groen HJM, Vermeulen KM, Greuter MJW, *et al.* Cost-effectiveness of lung cancer screening with low-dose computed tomography in heavy smokers: a microsimulation modelling study. *Eur J Cancer* 2020;**135**:121-9. https://dx.doi.org/10.1016/j.ejca.2020.05.004
125. 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**:1-146.
126. Goffin JR, Flanagan WM, Miller AB, Fitzgerald NR, Memon S, Wolfson MC, *et al.* Biennial lung cancer screening in Canada with smoking cessation-outcomes and cost-effectiveness. *Lung Cancer* 2016;**101**:98-103. https://dx.doi.org/10.1016/j.lungcan.2016.09.013

127. Goffin JR, Flanagan WM, Miller AB, Fitzgerald NR, Memon S, Wolfson MC, *et al.* Cost-effectiveness of lung cancer screening in Canada. *JAMA Oncol* 2015;**1**:807-13. https://dx.doi.org/10.1001/jamaoncol.2015.2472
128. Hinde S, Crilly T, Balata H, Bartlett R, Crilly J, Barber P, *et al.* The cost-effectiveness of the Manchester 'lung health checks', a community-based lung cancer low-dose CT screening pilot. *Lung Cancer* 2018;**126**:119-24. https://dx.doi.org/10.1016/j.lungcan.2018.10.029
129. Hofer F, Kauczor HU, Stargardt T. Cost-utility analysis of a potential lung cancer screening program for a high-risk population in Germany: a modelling approach. *Lung Cancer* 2018;**124**:189-98. https://dx.doi.org/10.1016/j.lungcan.2018.07.036
130. Jaine R, Kvizhinadze G, Nair N, Blakely T. Cost-effectiveness of a low-dose computed tomography screening programme for lung cancer in New Zealand. *Lung Cancer* 2018;**124**:233-40. https://dx.doi.org/10.1016/j.lungcan.2018.08.004
131. Mahadevia PJ, Fleisher LA, Frick KD, Eng J, Goodman SN, Powe NR. Lung cancer screening with helical computed tomography in older adult smokers: a decision and cost-effectiveness analysis. *Jama* 2003;**289**:313-22.
132. Manser R, Dalton A, Carter R, Byrnes G, Elwood M, Campbell DA. Cost-effectiveness analysis of screening for lung cancer with low dose spiral CT (computed tomography) in the Australian setting. *Lung Cancer* 2005;**48**:171-85.
133. Marshall D, Simpson KN, Earle CC, Chu CW. Economic decision analysis model of screening for lung cancer. *Eur J Cancer* 2001;**37**:1759-67.
134. Marshall D, Simpson KN, Earle CC, Chu CW. Potential cost-effectiveness of one-time screening for lung cancer (LC) in a high risk cohort. *Lung Cancer* 2001;**32**:227-36.
135. McLeod M, Sandiford P, Kvizhinadze G, Bartholomew K, Crengle S. Impact of low-dose CT screening for lung cancer on ethnic health inequities in New Zealand: a cost-effectiveness analysis. *BMJ Open* 2020;**10**:e037145. https://dx.doi.org/10.1136/bmjopen-2020-037145
136. Okamoto N. Cost-effectiveness of lung cancer screening in Japan. *Cancer* 2000;**89**:2489-93.
137. Pyenson BS, Henschke CI, Yankelevitz DF, Yip R, Dec E. Offering lung cancer screening to high-risk medicare beneficiaries saves lives and is cost-effective: an actuarial analysis. *Am Health Drug Benefits* 2014;**7**:272-82.
138. Pyenson BS, Sander MS, Jiang Y, Kahn H, Mulshine JL. An actuarial analysis shows that offering lung cancer screening as an insurance benefit would save lives at relatively low cost. *Health Affairs* 2012;**31**:770-9.
139. Shmueli A, Fraifeld S, Peretz T, Gutfeld O, Gips M, Sosna J, *et al.* Cost-effectiveness of baseline low-dose computed tomography screening for lung cancer: the Israeli experience. *Value Health* 2013;**16**:922-31.
140. Griffin E, Hyde C, Long L, Varley-Campbell J, Coelho H, Robinson S, *et al.* Lung cancer screening by low-dose computed tomography: a cost-effectiveness analysis of alternative programmes in the UK using a newly developed natural history-based economic model. *Diagn Progn Res* 2020;**4**:20. 10.1186/s41512-020-00087-y
141. 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**:1-276. https://dx.doi.org/10.3310/hta22690
142. Tabata H, Akita T, Matsuura A, Kaishima T, Matsuoka T, Ohisa M, *et al.* Cost-effectiveness of the introduction of low-dose CT screening in Japanese smokers aged 55 to 74 years old. *Hiroshima J Med Sci* 2014;**63**:13-22.
143. Criss SD, Cao P, Bastani M, Ten Haaf K, Chen Y, Sheehan DF, *et al.* Cost-effectiveness analysis of lung cancer screening in the United States: a comparative modeling study. *Ann Intern Med* 2019;**171**:796-804. https://dx.doi.org/10.7326/M19-0322
144. Tomonaga Y, Ten Haaf K, Frauenfelder T, Kohler M, Kouyos RD, Shilaih M, *et al.* Cost-effectiveness of low-dose CT screening for lung cancer in a European country with high prevalence of smoking - a modelling study. *Lung Cancer* 2018;**121**:61-9. https://dx.doi.org/10.1016/j.lungcan.2018.05.008
145. Ten Haaf K, Tammemagi MC, Bondy SJ, van der Aalst CM, Gu S, McGregor SE, *et al.* Performance and cost-effectiveness of computed tomography lung cancer screening scenarios in a

population-based setting: a microsimulation modeling analysis in Ontario, Canada. *PLoS Med* 2017;**14**:e1002225. https://dx.doi.org/10.1371/journal.pmed.1002225
146. Toumazis I, Tsai EB, Erdogan SA, Han SS, Wan W, Leung A*, et al.* Cost-effectiveness analysis of lung cancer screening accounting for the effect of indeterminate findings. *JNCI Cancer Spectr* 2019;**3**:pkz035. https://dx.doi.org/10.1093/jncics/pkz035
147. Treskova M, Aumann I, Golpon H, Vogel-Claussen J, Welte T, Kuhlmann A. Trade-off between benefits, harms and economic efficiency of low-dose CT lung cancer screening: a microsimulation analysis of nodule management strategies in a population-based setting. *BMC Med* 2017;**15**:162. https://dx.doi.org/10.1186/s12916-017-0924-3
148. Veronesi G, Navone N, Novellis P, Dieci E, Toschi L, Velutti L*, et al.* Favorable incremental cost-effectiveness ratio for lung cancer screening in Italy. *Lung Cancer* 2020;**143**:73-9. https://dx.doi.org/10.1016/j.lungcan.2020.03.015
149. Villanti AC, Jiang Y, Abrams DB, Pyenson BS. A cost-utility analysis of lung cancer screening and the additional benefits of incorporating smoking cessation interventions. *PLoS ONE* 2013;**8**:e71379. https://dx.doi.org/10.1371/journal.pone.0071379
150. Whynes DK. Could CT screening for lung cancer ever be cost effective in the United Kingdom? *Cost Effectiveness & Resource Allocation* 2008;**6**:5. https://dx.doi.org/10.1186/1478-7547-6-5
151. Wisnivesky JP, Mushlin AI, Sicherman N, Henschke C. The cost-effectiveness of low-dose CT screening for lung cancer: preliminary results of baseline screening. *Chest* 2003;**124**:614-21.
152. Yang SC, Lai WW, Lin CC, Su WC, Ku LJ, Hwang JS*, et al.* Cost-effectiveness of implementing computed tomography screening for lung cancer in Taiwan. *Lung Cancer* 2017;**108**:183-91. https://dx.doi.org/10.1016/j.lungcan.2017.04.001
153. McMahon PM, Kong CY, Bouzan C, Weinstein MC, Cipriano LE, Tramontano AC*, et al.* Cost-effectiveness of computed tomography screening for lung cancer in the United States. *J Thorac Oncol* 2011;**6**:1841-8.
154. McMahon PM, Kong CY, Johnson BE, Weinstein MC, Weeks JC, Tramontano AC*, et al.* Chapter 9: The MGH-HMS lung cancer policy model: tobacco control versus screening. *Risk Anal* 2012;**32**:S117-24. 10.1111/j.1539-6924.2011.01652.x
155. Hinde S, McKenna C, Whyte S, Peake MD, Callister ME, Rogers T*, et al.* Modelling the cost-effectiveness of public awareness campaigns for the early detection of non-small-cell lung cancer. *Br J Cancer* 2015;**113**:135-41. 10.1038/bjc.2015.167
156. Ten Haaf K, van Rosmalen J, de Koning HJ. Lung cancer detectability by test, histology, stage, and gender: estimates from the NLST and the PLCO trials. *Cancer Epidemiol Biomarkers Prev* 2015;**24**:154-61. 10.1158/1055-9965.Epi-14-0745
157. Infante M, Berghmans T, Heuvelmans MA, Hillerdal G, Oudkerk M. Slow-growing lung cancer as an emerging entity: from screening to clinical management. *Eur Respir J* 2013;**42**:1706-22. 10.1183/09031936.00186212
158. Tosi D, Nosotti M, Bonitta G, Mendogni P, Bertolaccini L, Spaggiari L*, et al.* Anatomical segmentectomy versus pulmonary lobectomy for stage I non-small-cell lung cancer: patients selection and outcomes from the European Society of Thoracic Surgeons database analysis. *Interactive CardioVascular and Thoracic Surgery* 2020;**32**:546-51. 10.1093/icvts/ivaa298
159. Jones GS, Baldwin DR. Recent advances in the management of lung cancer. *Clin Med (Lond)* 2018;**18**:s41-s6. 10.7861/clinmedicine.18-2-s41
160. National Institute for H, Care E. *Evidence standards framework for digital health technologies*. London: NICE; 2018.
161. Chapman C, Robertson J, Murray A, Titulaer M, Lang B, Thorpe A*, et al.* The presence of autoantibodies to tumour-associated antigens can predate clinical diagnosis of small cell lung cancer. *Chest* 2010;**138**:775A. http://dx.doi.org/10.1378/chest.10476
162. Chapman CJ, Murray A, Peek L, Boyle P, Allen J, Parsy-Kowalska C*, et al.* An improved EarlyCDT-Lung test can be used to aid the detection of lung cancer. *J Thorac Oncol* 2010;**5**:S548. http://dx.doi.org/10.1097/JTO.0b013e3182004f15
163. Holdenrieder S, Chapman CJ, Stieber P. Autoantibodies in comparison with cancer biomarkers in patients with advanced lung cancer. *Tumor Biol* 2011;**32**:S117. http://dx.doi.org/10.1007/s13277-011-0240-z

164. McElveen J, Allen J, Moorcorft L, Gilbert S, Hamilton-Fairley G, Murray A. Improvement in performance of the EarlyCDT-Lung test for detection of lung cancer by addition of antigenic biomarkers. *Tumour Biol* 2016;**37**:S21-2. http://dx.doi.org/10.1007/s13277-016-5287-4
165. Murray A, Chapman C, Parsons G, Peek L, Allen J, McElveen J*, et al.* Reproducibility of autoantibody measurements in normal individuals using the Early CDT-Lung test. *J Clin Oncol* 2010;**28**:e17522. http://dx.doi.org/10.1200/jco.2010.28.15_suppl.e17522
166. Peek L, Lam S, Healey G, Fritsche HA, Chapman C, Murray A*, et al.* Use of serum autoantibodies to identify early-stage lung cancer. *J Clin Oncol* 2010;**28**:7032.
167. Peek L, Murray A, Healey G, Jett J. Autoantibody test helps identify malignant pulmonary nodules. *J Thorac Oncol* 2018;**13**:S158-9. 10.1016/j.jtho.2018.07.035
168. Peking University People's Hospital. *A preliminary study on the detection of plasma markers in early diagnosis for lung cancer*. ClinicalTrials.gov. Bethesda (MD): National Library of Medicine (US); 2020. URL: https://ClinicalTrials.gov/show/NCT04558255 (accessed 9th March 2021).
169. Allen J, Healey GF, Parsy-Kowalska C, Macdonald IK, Hudson NJ, McElveen J*, et al.* Improvement in the performance of the EarlyCDT-lung test in the detection of lung cancer by the application of additional metrics to reduce false positive results. Paper presented at: American Thoracic Society 2015 International Conference; Denver, May 15-20.
170. Boyle P, Lam S, Maddison P, Wood WC, Peek LJ, Murray A*, et al.* Antigen positivity in the EarlyCDT-Lung test by stage and histological subtype of lung cancer. *Chest* 2010;**138**:727A. http://dx.doi.org/10.1378/chest.10504
171. Boyle P, Weycker D, Khuu A, Jett JR, Detterbeck FC, Kennedy TC*, et al.* An autoantibody test (AABT) to aid in early detection of lung cancer in high-risk patients is likely to be cost-effective. *Chest* 2010;**138**:241A. http://dx.doi.org/10.1378/chest.10709
172. Chang N, Zhang J, Yang T, Wu S. Evaluation of seven-autoantibody panel test in early detection of lung cancer in routine clinical practice. *J Thorac Oncol* 2019;**14**:S1010. 10.1016/j.jtho.2019.08.2230
173. Chapman C, Murray A, McElveen J, Sahin U, Luxemburger U, Tureci O*, et al.* Autoantibodies in lung cancer. *Tumour Biol* 2006;**27**:41.
174. Chapman C, Healey G, Murray A, Peek L, Robertson J. Autoantibodies in lung cancer and their use in early cancer diagnosis. *Tumour Biol* 2011;**32**:S31. http://dx.doi.org/10.1007/s13277-011-0238-6
175. Chapman C, Maddison P, Murray A, Thorpe A, Parsy-Kowalska C, Robertson J. Identification of tumor-associated autoantibodies in small cell lung cancer as immune markers of disease. *Chest* 2010;**138**:238A. http://dx.doi.org/10.1378/chest.10487
176. Chapman CJ, Hollevoet K, Van Meerbeeck J, Murray A, Peek L, Thorpe A*, et al.* Mesothelioma and tumor-associated autoantibodies. Paper presented at: American Thoracic Society 2011 International Conference; Denver, May 11-18.
177. Chapman C, McElveen J, Allen J, Gilbert S, Jett J, Mazzone P*, et al.* Improvement in performance of an autoantibody panel test for detection of lung cancer by addition of a single novel biomarker (EDB1). *J Thorac Oncol* 2017;**12**:S546.
178. Mayo Clinic. *Chest Computed Tomography (CT) screening study with antibody testing*. ClinicalTrials.gov. Bethesda (MD): National Library of Medicine (US); 2010. URL: https://ClinicalTrials.gov/show/NCT01203579 (accessed 9th March 2021).
179. Colpitts T, Russell J, Ramirez J, Russell E, Stroupe S. A panel of tumor antigens and autoantibodies for the detection of lung cancer. *Tumour Biol* 2007;**28**:91.
180. Du Q, Yu R, Wang H, Yan D, Yuan Q, Ma Y*, et al.* Significance of tumor-associated autoantibodies in the early diagnosis of lung cancer. *Clin Respir J* 2018;**12**:2020-8. https://dx.doi.org/10.1111/crj.12769
181. Eiermann W, Jackson L, Murray A, Chapman CJ, Peek LJ, Widschwendter P*, et al.* Serum autoantibodies to breast cancer associated antigens reflect tumor biology: an opportunity for early detection & prevention? *Cancer Res* 2011;**71**. http://dx.doi.org/10.1158/0008-5472.SABCS11-P4-08-03
182. Farlow EC, Kim AW, Liptay MJ, Patel K, Faber LP, Kaiser-Walters KA*, et al.* Tumor-specific auto-antibodies as an early detection screening tool for non-small cell lung cancer. *J Thorac Oncol* 2009;**4**:S612-S3.

183. Farlow EC, Patel K, Basu S, Lee BS, Kim AW, Coon JS*, et al.* Development of a multiplexed tumor-associated autoantibody-based blood test for the detection of non-small cell lung cancer. *Clin Cancer Res* 2010;**16**:3452-62. https://dx.doi.org/10.1158/1078-0432.CCR-09-3192
184. Farlow EC, Vercillo MS, Coon JS, Basu S, Kim AW, Faber LP*, et al.* A multi-analyte serum test for the detection of non-small cell lung cancer. *Br J Cancer* 2010;**103**:1221-8. 10.1038/sj.bjc.6605865
185. He Y, Zhou C, Hirsch F. Tumor autoantibody panel can improve the accuracy of early diagnosis in lung cancer presenting with GGNs /solid nodules. *J Thorac Oncol* 2018;**13**:S781-2. http://dx.doi.org/10.1016/j.jtho.2018.08.1354
186. Huang H, Luo W, Ni Y, Sun S, Wang C, Zhang L. The diagnostic efficiency of seven autoantibodies in lung cancer. *Eur J Cancer Prev* 2020;**29**:315-20. https://dx.doi.org/10.1097/CEJ.0000000000000559
187. Jamnani FR. The state of the art in the development of a panel of biomarkers for the early detection of lung cancer. *J Thorac Dis* 2018;**10**:625-7. 10.21037/jtd.2018.01.03
188. Jett J, Healey G, Macdonald I, Parsy-Kowalska C, Peek L, Murray A. Determination of the detection lead time for autoantibody biomarkers in early stage lung cancer using the UKCTOCS cohort. *J Thorac Oncol* 2017;**12**:S2170.
189. Jett J, Pitcher T, Doubleday K, Peek L. Identification of likely malignant indeterminate pulmonary nodules by analysis of autoantibodies against lung cancer-associated antigens. *Chest* 2020;**158**:A1464. http://dx.doi.org/10.1016/j.chest.2020.08.1321
190. Jia J, Wang W, Meng W, Ding M, Ma S, Wang X. Development of a multiplex autoantibody test for detection of lung cancer. *PLoS ONE* 2014;**9**:e95444. https://dx.doi.org/10.1371/journal.pone.0095444
191. Jia J, Huang J, Zhu X. [Diagnostic value of tumor-associated autoantibodies for malignant pulmonary nodules]. *Chinese Journal of Clinical Oncology* 2020;**47**:271-6. http://dx.doi.org/10.3969/j.issn.1000-8179.2020.06.272
192. Khattar NH, Coe-Atkinson SP, Stromberg AJ, Jett JR, Hirschowitz EA. Lung cancer-associated auto-antibodies measured using seven amino acid peptides in a diagnostic blood test for lung cancer. *Cancer Biol Ther* 2010;**10**:267-72.
193. Lastwika KJ, Kargl J, Zhang Y, Zhu X, Shelley D, Lo E*, et al.* Non-small cell lung tumor-derived autoantibodies can distinguish benign from malignant pulmonary nodules. *Clin Cancer Res* 2018;**24**:A12. http://dx.doi.org/10.1158/1557-3265.AACRIASLC18-A12
194. Lastwika KJ, Kargl J, Zhang Y, Zhu X, Lo E, Shelley D*, et al.* Tumor-derived autoantibodies identify malignant pulmonary nodules. *Am J Respir Crit Care Med* 2019;**199**:1257-66. https://dx.doi.org/10.1164/rccm.201804-0628OC
195. Lastwika KJ, Zhang Y, Ladd JJ, Massion PP, Houghton AM, Lampe PD. Autoantibody-antigen complexes can detect limited-stage small-cell lung cancer. *J Thorac Oncol* 2020;**15**:S17.
196. Lastwika KL, Zhang Y, Shipley M, Kinahan PE, Pipovath S, Wu V*, et al.* Plasma proteomic, glycomic, and autoantibody biomarkers for lung cancer early detection. *J Thorac Oncol* 2020;**15**:S3. http://dx.doi.org/10.1016/j.jtho.2019.12.008
197. Lu L, Ding M, Cheng Y, Guo Y, Wang K, Fan L. [The clinical application value of seven tumor-associated autoantibodies in the diagnosis of non-small cell lung cancer]. *Chinese Journal of Laboratory Medicine* 2019;**42**:768-75. http://dx.doi.org/10.3760/cma.j.issn.1009-9158.2019.09.010
198. Mathew J, Healey G, Jewell W, Murray A, Chapman C, Peek L*, et al.* Demographics of populations at high risk of lung cancer and results of the Early CDT-Lung test. *J Clin Oncol* 2010;**28**:7033.
199. Mathew J. *Spectrum of autoantibody response to tumour associated antigens in normal population.* Doctor of Medicine. Nottingham: University of Nottingham; 2013.
200. Mazzone P, Wang XF, Choi H, Seeley M, Doseeva V, Scherer R*, et al.* Discovery and validation of a combined serum protein and auto-antibody test for the identification of lung cancer in a high risk population. *Am J Respir Crit Care Med* 2016;**193**:A7537.
201. Mazzone P. Early detection biomarker development: is success on the horizon? *J Thorac Oncol* 2018;**13**:S302-3. http://dx.doi.org/10.1016/j.jtho.2018.08.200

202. Meng QC, Gao PR, Ren PF, Song YP, Li HL. [Early diagnosis of subtype in early clinical stage lung adenocarcinoma by using an autoantibody panel and computed tomography]. *Chinese Medical Journal* 2019;**99**:204-8. https://dx.doi.org/10.3760/cma.j.issn.0376-2491.2019.03.010
203. China-Japan Friendship Hospital. *Monitoring the changes of tumor-related biomarkers before and after pulmonary nodule biopsy*. ClinicalTrials.gov. Bethesda (MD): National Library of Medicine (US); 2018. URL: https://ClinicalTrials.gov/show/NCT03397355 (accessed 9th March 2021).
204. Mu Y, Xie F, Sun T. Clinical value of seven autoantibodies combined detection in the diagnosis of lung cancer. *J Clin Lab Anal* 2020;**34**:e23349. https://dx.doi.org/10.1002/jcla.23349
205. Murray A, Healey G, Chapman C, Parsons G, Peek L, Allen J*, et al.* Reproducibility of autoantibody measurements in normal individuals using the Early CDT-Lung test. *Chest* 2010;**138**:242A. http://dx.doi.org/10.1378/chest.10483
206. Murray A, Storr S, Bottomley L, McElveen J, Parsy-Kowalska C, Chapman CJ*, et al.* Demonstration of the specificity of human autoantibody responses to tumor-associated antigens. *Am J Respir Crit Care Med* 2011;**183**:A4388.
207. Johnson DW. *Non-invasive biomarkers for early detection of lung cancers (ISRUSAL01)*. ClinicalTrials.gov. Bethesda (MD): National Library of Medicine (US); 2012. URL: https://ClinicalTrials.gov/show/NCT01580332 (accessed 9th March 2021).
208. Pedchenko T, Mernaugh R, Parekh D, Li M, Massion PP. Early detection of NSCLC with scFv selected against IgM autoantibody. *PLoS ONE* 2013;**8**:e60934. https://dx.doi.org/10.1371/journal.pone.0060934
209. Peek LJ, Healey G, Murray A, Robertson JF. Early CDT-Lung: QC measures reveal high precision, accuracy, and robustness for a clinical test to aid in early lung cancer detection for high-risk individuals. *Chest* 2010;**138**:243A. http://dx.doi.org/10.1378/chest.10503
210. Ren S, Zhang S, Ma Z, Cai H, Xu X, Zhou J*, et al.* Validation of autoantibody panel for early detection of lung cancer in Chinese population. *J Thorac Oncol* 2015;**10**:S306.
211. Ren S, Zhang S, Ma Z, Cai HR, Xu X, Zhou J*, et al.* Validation of autoantibody panel for early detection of lung cancer in Chinese population. *J Clin Oncol* 2015;**33**:e22143.
212. Ren S, Zhang S, Jiang T, He Y, Ma Z, Cai H*, et al.* Early detection of lung cancer by using an autoantibody panel in Chinese population. *Oncoimmunology* 2018;**7**:e1384108. https://dx.doi.org/10.1080/2162402X.2017.1384108
213. Trudgen K, Khattar NH, Bensadoun E, Arnold S, Stromberg AJ, Hirschowitz EA. Autoantibody profiling for lung cancer screening longitudinal retrospective analysis of CT screening cohorts. *PLoS ONE* 2014;**9**:e87947. https://dx.doi.org/10.1371/journal.pone.0087947
214. Wang W, Zhuang R, Ma H, Fang L, Wang Z, Lv W*, et al.* The diagnostic value of a seven-autoantibody panel and a nomogram with a scoring table for predicting the risk of non-small-cell lung cancer. *Cancer Sci* 2020;**111**:1699-710. https://dx.doi.org/10.1111/cas.14371
215. Weycker D, Jett JR, Detterbeck FC, Miller DL, Khuu A, Kennedy TC*, et al.* Cost-effectiveness of an autoantibody test (AABT) as an aid to diagnosis of lung cancer. *J Clin Oncol* 2010;**28**:7030.
216. Weycker D, Boyle P, Lopez A, Jett JR, Detterbeck F, Kennedy TC*, et al.* Cost-effectiveness of screening older adult smokers for lung cancer with an autoantibody test (AABT). *Am J Respir Crit Care Med* 2011;**183**:A2937.
217. Yao Y, Wu Z, Zhou Q. [Autoantibodies as the early diagnostic biomarkers for lung cancer]. *Chinese Journal of Lung Cancer* 2010;**13**:903-7. https://dx.doi.org/10.3779/j.issn.1009-3419.2010.09.12
218. Yin-Yu M, Fu-Yi X, Fei-Bo W, Ri-Yi Z, Xiao-Lin Y. Performance evaluation of an enzyme-linked immunosorbent assay for seven autoantibodies in lung cancer. *Clin Lab* 2019;**65**. https://dx.doi.org/10.7754/Clin.Lab.2018.180930
219. Zhou C, Ren S, He Y. Diagnostic yield of autoantibody panel for patients with ground-glass nodules (GGNS) or solid nodules in Chinese population. *J Thorac Oncol* 2015;**10**:S615.
220. Gambhir SS, Shepherd JE, Shah BD, Hart E, Hoh CK, Valk PE*, et al.* Analytical decision model for the cost-effective management of solitary pulmonary nodules. *J Clin Oncol* 1998;**16**:2113-25. 10.1200/jco.1998.16.6.2113
221. Seely JM, Mayo JR, Miller RR, Müller NL. T1 lung cancer: prevalence of mediastinal nodal metastases and diagnostic accuracy of CT. *Radiology* 1993;**186**:129-32. 10.1148/radiology.186.1.8416552

222. Cummings SR, Lillington GA, Richard RJ. Managing solitary pulmonary nodules. The choice of strategy in a "close call". *Am Rev Respir Dis* 1986;**134**:453–60.
223. Swensen SJ, Jett JR, Hartman TE, Midthun DE, Sloan JA, Sykes AM*, et al.* Lung cancer screening with CT: Mayo Clinic experience. *Radiology* 2003;**226**:756-61. 10.1148/radiol.2263020036
224. Zwirewich CV, Vedal S, Miller RR, Müller NL. Solitary pulmonary nodule: high-resolution CT and radiologic-pathologic correlation. *Radiology* 1991;**179**:469-76. 10.1148/radiology.179.2.2014294
225. Swensen SJ, Brown LR, Colby TV, Weaver AL, Midthun DE. Lung nodule enhancement at CT: prospective findings. *Radiology* 1996;**201**:447-55. 10.1148/radiology.201.2.8888239
226. Midthun DE, Swensen SJ, Jett JR. Clinical strategies for solitary pulmonary nodule. *Annu Rev Med* 1992;**43**:195-208. 10.1146/annurev.me.43.020192.001211
227. Cummings SR, Lillington GA, Richard RJ. Estimating the probability of malignancy in solitary pulmonary nodules. A Bayesian approach. *Am Rev Respir Dis* 1986;**134**:449-52. 10.1164/arrd.1986.134.3.449
228. Office for National Statistics. *National life tables*. London: Office for National Statistics; 2018.
229. Sturza J. A review and meta-analysis of utility values for lung cancer. *Med Decis Making* 2010;**30**:685-93. 10.1177/0272989x10369004
230. Cancer Research UK. *Saving lives, averting costs: an analysis of the financial implications of achieving earlier diagnosis of colorectal, lung and ovarian cancer*. London: Cancer Research UK; 2014.
231. Vanni T, Karnon J, Madan J, White RG, Edmunds WJ, Foss AM*, et al.* Calibrating models in economic evaluation: a seven-step approach. *Pharmacoeconomics* 2011;**29**:35-49. 10.2165/11584600-000000000-00000
232. Lin RS, Plevritis SK. Comparing the benefits of screening for breast cancer and lung cancer using a novel natural history model. *Cancer Causes Control* 2012;**23**:175-85. 10.1007/s10552-011-9866-9
233. Krauth C. Health economic analysis of screening. *GMS Curr Top Otorhinolaryngol Head Neck Surg* 2008;**7**:1-14.
234. Mahnken JD, Chan W, Freeman DH, Jr., Freeman JL. Reducing the effects of lead-time bias, length bias and over-detection in evaluating screening mammography: a censored bivariate data approach. *Stat Methods Med Res* 2008;**17**:643-63. 10.1177/0962280207087309
235. Office for National Statistics. *Interim Life Tables: England and Wales, 2010–2012*. Newport: Office for National Statistics; 2013.
236. Office for National Statistics. *Deaths Registered in England and Wales, 2014: Table 5.2 Neoplasms.* Newport: Office for National Statistics; 2015.
237. Continuous Mortality Investigation Mortality Committee. *Working paper 21. The graduation of the CMI 1999–2002 mortality experience: final '00' series mortality tables – Assured lives*. London: Institute and Faculty of Actuaries; 2006.
238. Craig R, Mindell J. *Health survey for England 2014: health, social care and lifestyles. Summary of key findings*. Leeds: NHS Digital; 2014.
239. 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**:525-32. 10.3111/13696998.2015.1032974
240. Round J, Jones L, Morris S. Estimating the cost of caring for people with cancer at the end of life: a modelling study. *Palliat Med* 2015;**29**:899-907. 10.1177/0269216315595203
241. Kennedy MPT, Hall PS, Callister MEJ. Factors affecting hospital costs in lung cancer patients in the United Kingdom. *Lung Cancer* 2016;**97**:8-14. https://doi.org/10.1016/j.lungcan.2016.04.009
242. Government Actuary's Department. *Interim life tables methodology*. 2007. URL: https://webarchive.nationalarchives.gov.uk/20080609145324/http://www.gad.gov.uk/Demography_Data/Life_Tables/methodology.asp (accessed 22nd July 2021).
243. Walters S, Maringe C, Coleman MP, Peake MD, Butler J, Young N*, et al.* Lung cancer survival and stage at diagnosis in Australia, Canada, Denmark, Norway, Sweden and the UK: a population-based study, 2004–2007. *Thorax* 2013;**68**:551-64. 10.1136/thoraxjnl-2012-202297
244. Solomon T, Rachet B, Whitehead S, Coleman MP. *Cancer survival in England: patients diagnosed 2007–2011 and followed up to 2012*. London: Office for National Statistics; 2013.

245. Office for National Statistics. *Cancer survival by stage at diagnosis for England (experimental statistics): adults diagnosed 2012, 2013 and 2014 and followed up to 2015*. Newport: Office for National Statistics; 2016.
246. National Lung Cancer Audit Project Team. *National Lung Cancer Audit Report 2013*. Leeds: Health and Social Care Information Centre; 2013.

11 APPENDICES

11.1 Literature search strategies for EarlyCDT Lung studies

MEDLINE ALL

(includes: Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE Daily and Ovid MEDLINE)

via Ovid http://ovidsp.ovid.com/

Date range: 1946 to 5th March, 2021

Date searched: 8th March 2021

Records retrieved: 1323

1 EarlyCDT.af. (18)

2 Early CDT.af. (2)

3 Early-CDT.af. (2)

4 Early cancer detection test.af. (6)

5 or/1-4 (24)

6 ECLS trial$.af. (4)

7 5 or 6 (26)

8 Oncimmune.af. (31)

9 7 or 8 (48)

10 Autoantibodies/ (68737)

11 (autoantibod$ or auto-antibod$ or AABT or AAb or AAbs or TAAb or TAAbs).ti,ab. (57395)

12 10 or 11 (94716)

13 exp Lung Neoplasms/ (239412)

14 Solitary Pulmonary Nodule/ (4170)

15 ((lung$ or pulmonary or bronchial or bronchogenic) adj3 (neoplas$ or carcinoma$ or cancer$ or nodule$ or tumor$ or tumour$ or malign$ or adenocarcinoma$ or blastoma$)).ti,ab. (251893)

16 NSCLC.ti,ab. (45747)

17 SCLC.ti,ab. (8139)

18 ((lung$ or pulmonary) adj2 (lesion$ or mass or masses)).ti,ab. (16824)

19 ((noncalcified or non calcified) adj2 (nodule$ or lesion$ or mass or masses)).ti,ab. (454)

20 NCPN.ti,ab. (3)

21 ((ground-glass or solid or part-solid or subsolid or sub-solid) adj2 (nodule$ or lesion$ or mass or masses)).ti,ab. (9645)

22 ground glass opacit$.ti,ab. (4208)

23 (GGN or GGNs or GGO or GGOs).ti,ab. (1546)

24 ((benign or malignant or indeterminate) adj2 nodule$).ti,ab. (5427)

25 coin lesion$.ti,ab. (484)

26 (IPN or IPNs).ti,ab. (1722)

27 or/13-26 (356910)

28 12 and 27 (1318)

29 9 or 28 (1345)

30 exp animals/ not humans.sh. (4796559)

31 29 not 30 (1323)

Key:

/ = subject heading (MeSH heading)

sh = subject heading (MeSH heading)

exp = exploded subject heading (MeSH heading)

$ = truncation

af = search of all fields

ti,ab = terms in title or abstract fields

adj3 = terms within three words of each other (any order)

Embase

via Ovid http://ovidsp.ovid.com/

Date range: 1974 to 5th March 2021

Date searched: 8th March 2021

Records retrieved: 1973

1 EarlyCDT.af. (56)

2 Early CDT.af. (17)

3 Early-CDT.af. (17)

4 Early cancer detection test.af. (14)

5 or/1-4 (71)

6 ECLS trial$.af. (9)

7 Oncimmune.af. (81)

8 5 or 6 or 7 (121)

9 autoantibody/ (73725)

10 (autoantibod$ or auto-antibod$ or AABT or AAb or AAbs or TAAb or TAAbs).ti,ab. (83322)

11 9 or 10 (104669)

12 exp lung tumor/ (391077)

13 lung nodule/ (22179)

14 lung coin lesion/ (560)

15 ((lung$ or pulmonary or bronchial or bronchogenic) adj3 (neoplas$ or carcinoma$ or cancer$ or nodule$ or tumor$ or tumour$ or malign$ or adenocarcinoma$ or blastoma$)).ti,ab. (364778)

16 NSCLC.ti,ab. (86396)

17 SCLC.ti,ab. (13679)

18 ((lung$ or pulmonary) adj2 (lesion$ or mass or masses)).ti,ab. (24742)

19 ((noncalcified or non calcified) adj2 (nodule$ or lesion$ or mass or masses)).ti,ab. (715)

20 NCPN.ti,ab. (4)

21 ((ground-glass or solid or part-solid or subsolid or sub-solid) adj2 (nodule$ or lesion$ or mass or masses)).ti,ab. (15482)

22 ground glass opacit$.ti,ab. (7717)

23 (GGN or GGNs or GGO or GGOs).ti,ab. (2499)

24 ((benign or malignant or indeterminate) adj2 nodule$).ti,ab. (8324)

25 coin lesion$.ti,ab. (460)

26 (IPN or IPNs).ti,ab. (2268)

27 or/12-24 (540557)

28 11 and 27 (1968)

29 8 or 28 (2021)

30 (animal/ or animal experiment/ or animal model/ or animal tissue/ or nonhuman/) not exp human/ (6204998)

31 29 not 30 (1973)

Key:

/ = subject heading (Emtree heading)

sh = subject heading (Emtree heading)

exp = exploded subject heading (Emtree heading)

$ = truncation

af = search of all fields

ti,ab = terms in title or abstract fields

adj3 = terms within three words of each other (any order)

Cochrane Central Register of Controlled Trials (CENTRAL)

via Wiley http://onlinelibrary.wiley.com/

Date range: Issue 3 of 12, March 2021

Date searched: 8th March 2021

Records retrieved: 29

The following strategy was used to search both CENTRAL and CDSR.

#1 EarlyCDT:ti,ab,kw 5

#2 "Early CDT":ti,ab,kw 13

#3 Early-CDT:ti,ab,kw 13

#4 "Early Cancer Detection Test":ti,ab,kw 4

#5 (OR #1-#4) 14

#6 ECLS next trial*:ti,ab,kw 3

#7 Oncimmune:ti,ab,kw 0

#8 #5 or #6 or #7 14

#9 MeSH descriptor: [Autoantibodies] this term only 686

#10 (autoantibod* or auto next antibod* or AABT or AAb or AAbs or TAAb or TAAbs):ti,ab,kw 1932

#11 #9 or #10 1932

#12 MeSH descriptor: [Lung Neoplasms] explode all trees 7828

#13 MeSH descriptor: [Solitary Pulmonary Nodule] this term only 81

#14 ((lung* or pulmonary or bronchial or bronchogenic) near/3 (neoplas* or carcinoma* or cancer* or nodule* or tumor* or tumour* or malign* or adenocarcinoma* or blastoma*)):ti,ab,kw 23544

#15 NSCLC:ti,ab,kw 9385

#16 SCLC:ti,ab,kw 1351

#17 ((lung* or pulmonary) near/2 (lesion* or mass or masses)):ti,ab,kw 600

#18 ((noncalcified or non calcified) near/2 (nodule* or lesion* or mass or masses)):ti,ab,kw 1478

#19 NCPN:ti,ab,kw 0

#20 ((ground next glass or solid or part next solid or subsolid or sub next solid) near/2 (nodule* or lesion* or mass or masses)):ti,ab,kw 469

#21 ground next glass next opacit*:ti,ab,kw 120

#22 (GGN or GGNs or GGO or GGOs):ti,ab,kw 67

#23 ((benign or malignant or indeterminate) near/2 nodule*):ti,ab,kw 226

#24 coin next lesion*:ti,ab,kw 3

#25 (IPN or IPNs):ti,ab,kw 38

#26 (OR #12-#25) 26367

#27 #11 AND #26 24

#28 #8 or #27 29

Key:

MeSH descriptor = subject heading (MeSH heading)

* = truncation

ti,ab,kw = terms in title, abstract or keyword fields

near/3 = terms within three words of each other (any order)

next = terms are next to each other

Science Citation Index

via Web of Science, Clarivate Analytics https://clarivate.com/

Date range: 1900 – 5th March 2021

Date searched: 8th March 2021

Records retrieved: 1536

# 22	1,536	#21 OR #7
# 21	1,513	#20 AND #8
# 20	387,884	#19 OR #18 OR #17 OR #16 OR #15 OR #14 OR #13 OR #12 OR #11 OR #10 OR #9
# 19	5,619	TS=(IPN or IPNs)
# 18	246	TS=("coin lesion" or "coin lesions")
# 17	5,559	TS=((benign or malignant or indeterminate) NEAR/2 nodule*)
# 16	1,516	TS=(GGN or GGNs or GGO or GGOs)
# 15	3,427	TS=("ground glass opacity" or "ground glass opacities")
# 14	12,363	TS=((ground-glass or solid or part-solid or subsolid or sub-solid) NEAR/2 (nodule* or lesion* or mass or masses))
# 13	7	TS=NCPN
# 12	439	TS=((noncalcified or non-calcified) NEAR/2 (nodule* or lesion* or mass or masses))
# 11	16,328	TS=((lung* or pulmonary) NEAR/2 (lesion* or mass or masses))
# 10	68,356	TS=(NSCLC or SCLC)
# 9	344,322	TS=((lung* or pulmonary or bronchial or bronchogenic) NEAR/3 (neoplas* or carcinoma* or cancer* or nodule* or tumor* or tumour* or malign* or adenocarcinoma* or blastoma*))
# 8	77,236	TS=(autoantibod* or auto-antibod* or AABT or AAb or AAbs or TAAb or TAAbs)
# 7	45	#6 OR #5
# 6	2	TS=Oncimmune
# 5	45	#4 OR #3
# 4	5	TS=("ECLS trial" or "ECLS trials")
# 3	41	#2 OR #1
# 2	8	TS=("Early cancer detection test")
# 1	35	TS=(EarlyCDT or "Early CDT" or Early-CDT)

Key:

TS = topic tag; searches in title, abstract, author keywords and keywords plus fields

* = truncation

NEAR/3 = terms within three words of each other (any order)

EconLit

via Ovid http://ovidsp.ovid.com/

Date range: 1886 to February 18, 2021

Date searched: 8th March 2021

Records retrieved: 3

1 EarlyCDT.af. (0)

2 Early CDT.af. (0)

3 Early-CDT.af. (0)

4 Early cancer detection test.af. (0)

5 1 or 2 or 3 or 4 (0)

6 ECLS trial$.af. (0)

7 Oncimmune.af. (0)

8 5 or 6 or 7 (0)

9 (autoantibod$ or auto-antibod$ or AABT or AAb or AAbs or TAAb or TAAbs).mp. (3)

Key:

$ = truncation

af = search of all fields

mp = terms in title, abstract, keywords, subject heading fields

Cochrane Database of Systematic Reviews (CDSR)

via Wiley http://onlinelibrary.wiley.com/

Date range: Issue 3 of 12, March 2021

Date searched: 8th March 2021

Records retrieved: 0

See above under CENTRAL for search strategy used.

Database of Abstracts of Reviews of Effects (DARE)

via http://www.crd.york.ac.uk/CRDWeb/

Date range: Inception – 31st March 2015

Date searched: 8th March 2021

Records retrieved: 1

The following strategy was used to search all CRD databases: DARE, HTA Database and NHS EED.

1 (EarlyCDT) OR (Early-CDT) OR ("Early CDT") 1

2	("Early Cancer Detection Test")	0
3	("ECLS trial") OR ("ECLS trials")	0
4	(Oncimmune)	1
5	#1 OR #2 OR #3 OR #4	1
6	(MeSH DESCRIPTOR Autoantibodies)	46
7	(autoantibod* or auto-antibod* or AABT or AAb or AAbs or TAAb or TAAbs)	67
8	#6 OR #7	67
9	(MeSH DESCRIPTOR Lung Neoplasms EXPLODE ALL TREES)	1151
10	(MeSH DESCRIPTOR Solitary Pulmonary Nodule)	27
11	((lung* or pulmonary or bronchial or bronchogenic) adj3 (neoplas* or carcinoma* or cancer* or nodule* or tumor* or tumour* or malign* or adenocarcinoma* or blastoma*))	1449
12	((neoplas* or carcinoma* or cancer* or nodule* or tumor* or tumour* or malign* or adenocarcinoma* or blastoma*) adj3 (lung* or pulmonary or bronchial or bronchogenic))	891
13	(NSCLC or SCLC)	284
14	((lung* or pulmonary) adj2 (lesion* or mass or masses)) OR ((lesion* or mass or masses) adj2 (lung* or pulmonary))	64
15	((noncalcified or non-calcified) adj2 (nodule* or lesion* or mass or masses)) OR ((nodule* or lesion* or mass or masses) adj2 (noncalcified or non-calcified))	6
16	(NCPN)	0
17	((ground-glass or solid or part-solid or subsolid or sub-solid) adj2 (nodule* or lesion* or mass or masses))	24
18	((nodule* or lesion* or mass or masses) adj2 (ground-glass or solid or part-solid or subsolid or sub-solid))	1
19	("ground glass opacity" OR "ground glass opacities")	2
20	(GGN or GGNs or GGO or GGOs)	0
21	((benign or malignant or indeterminate) adj2 nodule*) OR (nodule* adj2 (benign or malignant or indeterminate))	53
22	("coin lesion") OR ("coin lesions")	1
23	(IPN or IPNs)	0
24	#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 OR #22 OR #23	1565
25	#8 AND #24	1
26	#5 OR #25	2

Health Technology Assessment (HTA) Database

via http://www.crd.york.ac.uk/CRDWeb/

Date range: Inception – 31st March 2018

Date searched: 8th March 2021

Records retrieved: 1

See above under DARE for search strategy used.

NHS Economic Evaluation Database (NHS EED)

via http://www.crd.york.ac.uk/CRDWeb/

Date range: Inception – 31st March 2015

Date searched: 8th March 2021

Records retrieved: 0

See above under DARE for search strategy used.

International Health Technology Assessment Database

via https://database.inahta.org/

Date range: Inception – 9th March 2021

Date searched: 10th March 2021

Records retrieved: 16

((((IPN or IPNs)[Title] OR (IPN or IPNs)[abs] OR (IPN or IPNs)[Keywords]) OR (("coin lesions")[Title] OR ("coin lesions")[abs] OR ("coin lesions")[Keywords]) OR (("coin lesion")[Title] OR ("coin lesion")[abs] OR ("coin lesion")[Keywords]) OR (((nodule*)[Title] OR (nodule*)[abs] OR (nodule*)[Keywords]) AND ((benign or malignant or indeterminate)[Title] OR (benign or malignant or indeterminate)[abs] OR (benign or malignant or indeterminate)[Keywords])) OR ((GGN or GGNs or GGO or GGOs)[Title] OR (GGN or GGNs or GGO or GGOs)[abs] OR (GGN or GGNs or GGO or GGOs)[Keywords]) OR (("ground glass opacities")[Title] OR ("ground glass opacities")[abs] OR ("ground glass opacities")[Keywords]) OR (("ground glass opacity")[Title] OR ("ground glass opacity")[abs] OR ("ground glass opacity")[Keywords]) OR ((("ground-glass" or "ground glass" or solid or "part-solid" or "part solid" or subsolid or "sub-solid" or "sub solid")[Title] OR ("ground-glass" or "ground glass" or solid or "part-solid" or "part solid" or subsolid or "sub-solid" or "sub solid")[abs] OR ("ground-glass" or "ground glass" or solid or "part-solid" or "part solid" or subsolid or "sub-solid" or "sub solid")[Keywords]) AND ((nodule* or lesion* or mass or masses)[Title] OR (nodule* or lesion* or mass or masses)[abs] OR (nodule* or lesion* or mass or masses)[Keywords])) OR ((NCPN)[Title] OR (NCPN)[abs] OR (NCPN)[Keywords]) OR (((nodule* or lesion* or mass or masses)[Title] OR (nodule* or lesion* or mass or masses)[abs] OR (nodule* or lesion* or mass or masses)[Keywords]) AND ((noncalcified or "non calcified" or "non-calcified")[Title] OR (noncalcified or "non calcified" or "non-calcified")[abs] OR (noncalcified or "non calcified" or "non-calcified")[Keywords])) OR (((lesion* or mass or masses)[Title] OR (lesion* or mass or masses)[abs] OR (lesion* or mass or masses)[Keywords]) AND ((lung* or pulmonary)[Title] OR (lung* or pulmonary)[abs] OR (lung* or pulmonary)[Keywords])) OR ((NSCLC or SCLC)[Title] OR (NSCLC or SCLC)[abs] OR (NSCLC or SCLC)[Keywords]) OR (((neoplas* or carcinoma* or cancer* or nodule* or tumor* or tumour* or malign* or adenocarcinoma* or blastoma*)[Title] OR (neoplas* or carcinoma* or cancer* or nodule* or tumor* or tumour* or malign* or adenocarcinoma* or blastoma*)[abs] OR (neoplas* or carcinoma* or cancer* or nodule* or tumor* or tumour* or malign* or adenocarcinoma* or blastoma*)[Keywords]) AND ((lung* or pulmonary or bronchial or bronchogenic)[Title] OR (lung* or pulmonary or bronchial or bronchogenic)[abs] OR (lung* or pulmonary or bronchial or bronchogenic)[Keywords])) OR ("Solitary Pulmonary Nodule"[mh]) OR ("Lung Neoplasms"[mhe])) AND (((autoantibod* or auto-antibod* or AABT or AAb or AAbs or TAAb or TAAbs)[Title] OR (autoantibod* or auto-antibod* or AABT or AAb or AAbs or TAAb or

TAAbs)[abs] OR (autoantibod* or auto-antibod* or AABT or AAb or AAbs or TAAb or TAAbs)[Keywords]) OR ("Autoantibodies"[mh]))) OR (((Oncimmune)[Title] OR (Oncimmune)[abs] OR (Oncimmune)[Keywords]) OR ((“ECLS trials”)[Title] OR (“ECLS trials”)[abs] OR (“ECLS trials”)[Keywords]) OR ((“ECLS trial”)[Title] OR (“ECLS trial”)[abs] OR (“ECLS trial”)[Keywords]) OR (((“Early Cancer Detection Test”)[Title] OR (“Early Cancer Detection Test”)[abs] OR (“Early Cancer Detection Test”)[Keywords]) AND ((lung* or pulmonary)[Title] OR (lung* or pulmonary)[abs] OR (lung* or pulmonary)[Keywords])) OR (("Early-CDT")[Title] OR ("Early-CDT")[abs] OR ("Early-CDT")[Keywords]) OR ((“Early CDT”)[Title] OR (“Early CDT”)[abs] OR (“Early CDT”)[Keywords]) OR ((EarlyCDT)[Title] OR (EarlyCDT)[abs] OR (EarlyCDT)[Keywords]))

Key:

[Keywords] = search of keywords field

[abs] = search of abstract field

[Title] = search of title field

[mh] = subject heading search

[mhe] = exploded subject heading search

* = truncation

ClinicalTrials.gov

https://clinicaltrials.gov/

Searched on: 9th March 2021

Records retrieved: 27

Advanced search used.

1. 4 Studies found for: EarlyCDT OR Early-CDT OR "Early CDT"

2. 3 Studies found for: "Early Cancer Detection Test"

3. 11 Studies found for: (autoantibody OR auto-antibody OR AABT OR AAb OR AAbs OR TAAb OR TAAbs) | lung cancer

4. 3 Studies found for: (autoantibody OR auto-antibody OR AABT OR AAb OR AAbs OR TAAb OR TAAbs) | pulmonary nodule

5. 3 Studies found for: (autoantibody OR auto-antibody OR AABT OR AAb OR AAbs OR TAAb OR TAAbs) | NSCLC OR SCLC

6. 1 Study found for: (autoantibody OR auto-antibody OR AABT OR AAb OR AAbs OR TAAb OR TAAbs) | coin lesion

7. 1 Study found for: (autoantibody OR auto-antibody OR AABT OR AAb OR AAbs OR TAAb OR TAAbs) | (indeterminate nodule OR IPN OR IPNs)

8. No Studies found for: (autoantibody OR auto-antibody OR AABT OR AAb OR AAbs OR TAAb OR TAAbs) | (ground glass OR GGN OR GGNs OR GGO OR GGOs)

9. No Studies found for: (autoantibody OR auto-antibody OR AABT OR AAb OR AAbs OR TAAb OR TAAbs) | NCPN OR noncalcified OR non-calcified)

10. 1 Study found for: Oncimmune

Key:

| = combine with AND

EU Clinical Trials Register

https://www.clinicaltrialsregister.eu/ctr-search/search

Searched on: 9th March 2021

Records retrieved: 8

1. EarlyCDT OR Early-CDT OR "Early CDT" – 0 results

2. "Early Cancer Detection Test" – 0 results

3. “ECLS trial OR “ECLS trials” – 0 results

4. Oncimmune – 0 results

5. 3 result(s) found for: (autoantibody OR auto-antibody OR autoantibodies OR auto-antibodies) AND (Lung OR pulmonary) AND (neoplasm OR carcinoma OR cancer OR nodule OR tumor OR tumour)

6. 2 result(s) found for: (autoantibody OR auto-antibody OR autoantibodies OR auto-antibodies) AND (NSCLC OR SCLC)

7. 3 result(s) found for: (autoantibody OR auto-antibody OR autoantibodies OR auto-antibodies) AND (Lung OR pulmonary) AND (lesion OR lesions OR mass OR masses)

8. (autoantibody OR auto-antibody OR autoantibodies OR auto-antibodies) AND (noncalcified or “non calcified” OR non-calcified) AND (nodule OR lesion OR lesions OR mass OR masses) – 0 results

9. (autoantibody OR auto-antibody OR autoantibodies OR auto-antibodies) AND NCPN – 0 results

10. (autoantibody OR auto-antibody OR autoantibodies OR auto-antibodies) AND (ground-glass or “ground glass” OR solid OR part-solid OR “part solid” OR subsolid OR sub-solid OR “sub solid”) – 0 results

11. (autoantibody OR auto-antibody OR autoantibodies OR auto-antibodies) AND (GGN OR GGNs OR GGO OR GGOs) – 0 results

12. (autoantibody OR auto-antibody OR autoantibodies OR auto-antibodies) AND (benign OR malignant OR indeterminate) AND (nodule OR nodules) – 0 results

13. (autoantibody OR auto-antibody OR autoantibodies OR auto-antibodies) AND (benign OR malignant OR indeterminate) AND (“coin lesion” OR “coin lesions”) – 0 results

14. (autoantibody OR auto-antibody OR autoantibodies OR auto-antibodies) AND (IPN OR IPNs) – 0 results

Conference Proceedings Citation Index: Science

via Web of Science, Clarivate Analytics https://clarivate.com/

Date range: 1990 – 5th March 2021

Date searched: 8th March 2021

Records retrieved: 75

# 22	75	#21 OR #7
# 21	69	#20 AND #8
# 20	51,033	#19 OR #18 OR #17 OR #16 OR #15 OR #14 OR #13 OR #12 OR #11 OR #10 OR #9
# 19	832	TS=(IPN or IPNs)
# 18	12	TS=("coin lesion" or "coin lesions")
# 17	657	TS=((benign or malignant or indeterminate) NEAR/2 nodule*)
# 16	182	TS=(GGN or GGNs or GGO or GGOs)
# 15	208	TS=("ground glass opacity" or "ground glass opacities")
# 14	1,403	TS=((ground-glass or solid or part-solid or subsolid or sub-solid) NEAR/2 (nodule* or lesion* or mass or masses))
# 13	1	TS=NCPN
# 12	39	TS=((noncalcified or non-calcified) NEAR/2 (nodule* or lesion* or mass or masses))
# 11	1,461	TS=((lung* or pulmonary) NEAR/2 (lesion* or mass or masses))
# 10	10,680	TS=(NSCLC or SCLC)
# 9	43,692	TS=((lung* or pulmonary or bronchial or bronchogenic) NEAR/3 (neoplas* or carcinoma* or cancer* or nodule* or tumor* or tumour* or malign* or adenocarcinoma* or blastoma*))
# 8	6,780	TS=(autoantibod* or auto-antibod* or AABT or AAb or AAbs or TAAb or TAAbs)
# 7	7	#6 OR #5
# 6	0	TS=Oncimmune
# 5	7	#4 OR #3
# 4	0	TS=("ECLS trial" or "ECLS trials")
# 3	7	#2 OR #1
# 2	3	TS=(“Early cancer detection test”)
# 1	4	TS=(EarlyCDT or "Early CDT" or Early-CDT)

Key:

TS = topic tag; searches in title, abstract, author keywords and keywords plus fields

* = truncation

NEAR/3 = terms within three words of each other (any order)

Proquest Dissertations & Theses A&I

via Proquest https://www.proquest.com/

Searched on: 9th March 2021

Records retrieved: 28

(TI,AB,SU,IF(EarlyCDT OR "Early CDT" OR Early-CDT) OR TI,AB,SU,IF("Early Cancer Detection Test") OR TI,AB,SU,IF("ECLS trial" OR "ECLS trials") OR TI,AB,SU,IF(Oncimmune)) OR (TI,AB,SU,IF(autoantibod* OR auto-antibod* OR AABT OR AAb OR AAbs OR TAAb OR TAAbs) AND (TI,AB,SU,IF((lung* OR pulmonary OR bronchial OR bronchogenic) NEAR/3 (neoplas* OR carcinoma* OR cancer* OR nodule* OR tumor* OR tumour* OR malign* OR adenocarcinoma* OR blastoma*)) OR TI,AB,SU,IF(NSCLC OR SCLC) OR TI,AB,SU,IF((lung* OR pulmonary) NEAR/2 (lesion* OR mass OR masses)) OR TI,AB,SU,IF((noncalcified OR non-calcified) NEAR/2 (nodule* OR lesion* OR mass OR masses)) OR TI,AB,SU,IF(NCPN) OR TI,AB,SU,IF((ground-glass OR solid OR part-solid OR subsolid OR sub-solid) NEAR/2 (nodule* OR lesion* OR mass OR masses)) OR TI,AB,SU,IF("ground glass opacity" OR "ground glass opacities") OR TI,AB,SU,IF(GGN OR GGNs OR GGO OR GGOs) OR TI,AB,SU,IF((benign OR malignant OR indeterminate) NEAR/2 nodule*) OR TI,AB,SU,IF("coin lesion" OR "coin lesions") OR TI,AB,SU,IF(IPN OR IPNs)))

Key:

TI,AB,SU,IF = search of title, abstract, subject headings, keyword fields

* = truncation

Open Access Theses and Dissertations

https://oatd.org/

Searched on: 9th March 2021

Records retrieved: 52

1. “EarlyCDT Lung" OR "Early-CDT Lung" OR "Early CDT Lung" 1 hit

2. Oncimmune – 1 hit

3. (autoantibod* OR auto-antibod* OR "auto antibody" OR "auto antibodies") AND (lung* OR pulmonary) AND (neoplas* OR carcinoma* OR cancer* OR nodule* OR tumor* OR tumour* OR lesion OR mass OR masses OR indeterminate OR "coin lesion" OR "coin lesions" OR "ground glass" OR ground-glass) 35 hits

4. (autoantibod* OR auto-antibod* OR "auto antibody" OR "auto antibodies") AND (NSCLC OR SCLC) 8 hits

5. (autoantibod* OR auto-antibod* OR "auto antibody" OR "auto antibodies") AND (noncalcified OR "non calcified" OR non-calcified) AND (nodule* or lesion* or mass or masses) AND (lung OR pulmonary) 7 hits

Key:

* = truncation

International Prospective Register of Systematic Reviews (PROSPERO)

via https://www.crd.york.ac.uk/prospero/

Date range: Inception – 8th March 2021

Date searched: 9th March 2021

Records retrieved: 0

#1 EarlyCDT or "Early CDT" or Early-CDT 0

#2 "Early Cancer Detection Test" 0

#3 "ECLS" trial or "ECLS trials" 0

#4 Oncimmune 0

#5 MeSH DESCRIPTOR Autoantibodies 20

#6 autoantibod* or auto-antibod* or AABT or AAb or AAbs or TAAb or TAAbs 195

#7 #5 OR #6 199

#8 MeSH DESCRIPTOR Lung Neoplasms EXPLODE ALL TREES 475

#9 MeSH DESCRIPTOR Solitary Pulmonary Nodule 5

#10 (lung* or pulmonary or bronchial or bronchogenic) adj3 (neoplas* or carcinoma* or cancer* or nodule* or tumor* or tumour* or malign* or adenocarcinoma* or blastoma*) 1337

#11 (neoplas* or carcinoma* or cancer* or nodule* or tumor* or tumour* or malign* or adenocarcinoma* or blastoma*) adj3 (lung* or pulmonary or bronchial or bronchogenic) 606

#12 NSCLC or SCLC 502

#13 (lung* or pulmonary) adj2 (lesion* or mass or masses) 47

#14 (lesion* or mass or masses) adj2 (lung* or pulmonary) 31

#15 (noncalcified or "non calcified" or non-calcified) adj2 (nodule* or lesion* or mass or masses) 1

#16 (nodule* or lesion* or mass or masses) adj2 (noncalcified or "non calcified" or non-calcified) 0

#17 NCPN 0

#18 (ground-glass or solid or part-solid or subsolid or sub-solid) adj2 (nodule* or lesion* or mass or masses) 41

#19 ("ground glass" or "part solid" or "sub solid") adj2 (nodule* or lesion* or mass or masses) 10

#20 (nodule* or lesion* or mass or masses) adj2 (ground-glass or solid or part-solid or subsolid or sub-solid) 13

#21 (nodule* or lesion* or mass or masses) adj2 ("ground glass" or "part solid" or "sub solid") 2

#22 "ground glass opacity" or "ground glass opacities" 40

#23 GGN or GGNs or GGO or GGOs 22

#24 (benign or malignant or indeterminate) adj2 nodule* 44

#25 nodule* adj2 (benign or malignant or indeterminate) 16

#26 "coin lesion" or "coin lesions" 1

#27 IPN or IPNs 8

#28 #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 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 1636

#29 #7 AND #28 0

Key:

MeSH DESCRIPTOR = subject heading (MeSH heading)

* = truncation

adj3 = terms within 3 words of each other (order specified)

11.2 Quality and risk of bias assessment

Study details

Reference	Sullivan et al 2021: "Earlier diagnosis of lung cancer in a randomised trial of an autoantibody blood test followed by imaging"

Study design

X Individually-randomized parallel-group trial

☐ Cluster-randomized parallel-group trial

☐ Individually randomized cross-over (or other matched) trial

For the purposes of this assessment, the interventions being compared are defined as

Experimental:	EarlyCDT Lung test	Comparator:	Standard clinical care

Specify which outcome is being assessed for risk of bias	Rate of stage III/IV lung cancer within 2 years of randomisation
Specify the numerical result being assessed. In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.	HR 0.64 (95% CI 0.41–0.99)

Is the review team's aim for this result…?

x to assess the effect of *assignment to intervention* (the 'intention-to-treat' effect)

☐ to assess the effect of *adhering to intervention* (the 'per-protocol' effect)

If the aim is to assess the effect of *adhering to intervention*, select the deviations from intended intervention that should be addressed (at least one must be checked):

☐ occurrence of non-protocol interventions

☐ failures in implementing the intervention that could have affected the outcome

☐ non-adherence to their assigned intervention by trial participants

Which of the following sources were <u>obtained</u> to help inform the risk-of-bias assessment? (tick as many as apply)

x	Journal article(s) with results of the trial
x	Trial protocol
☐	Statistical analysis plan (SAP)
x	Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
☐	Company-owned trial registry record (e.g. GSK Clinical Study Register record)
☐	"Grey literature" (e.g. unpublished thesis)
x	Conference abstract(s) about the trial
☐	Regulatory document (e.g. Clinical Study Report, Drug Approval Package)
☐	Research ethics application
☐	Grant database summary (e.g. NIH RePORTER or Research Councils UK Gateway to Research)
☐	Personal communication with trialist
☐	Personal communication with the sponsor

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Domain 1: Risk of bias arising from the randomization process

Signalling questions	Comments	Response options
1.1 Was the allocation sequence random?	Used a web-based randomisation system provided by Tayside Clinical Trials Unit	Y
1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?		Y
1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	Randomisation procedure used minimisation to ensure age, sex and smoking history were balanced across groups	N
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias arising from the randomization process?		NA

Domain 2: Risk of bias due to deviations from the intended interventions (*effect of assignment to intervention*)

Signalling questions	Comments	Response options
2.1. Were participants aware of their assigned intervention during the trial?	Blinding was not possible	Y
2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?		Y
2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the trial context?		PN
2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?		NA
2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?		NA
2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?		Y
2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?		NA
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to deviations from intended interventions?		NA

Domain 3: Missing outcome data

Signalling questions	Comments	Response options
3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Only one patient with missing data	Y
3.2 If N/PN/NI to 3.1: Is there evidence that the result was not biased by missing outcome data?		NA
3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?		NA
3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?		NA
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to missing outcome data?		NA

Domain 4: Risk of bias in measurement of the outcome

Signalling questions	Comments	Response options
4.1 Was the method of measuring the outcome inappropriate?		N
4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?		N
4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants?	Pathology and tumour staging reports were prepared by independent assessors who were blinded to the allocation status of participants. Staging data were taken from the Scottish Cancer Registry.	N
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?		NA
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?		NA
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias in measurement of the outcome?		NA

Domain 5: Risk of bias in selection of the reported result

Signalling questions	Comments	Response options
5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Statistical analysis plan available as supplementary file to the main published paper.	Y
Is the numerical result being assessed likely to have been selected, on the basis of the results, from...		
5.2. ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?		N
5.3 ... multiple eligible analyses of the data?		N
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to selection of the reported result?		NA

Overall risk of bias

Risk-of-bias judgement		Low
Optional: What is the overall predicted direction of bias for this outcome?		NA

11.3 Data abstraction tables

Table 29 presents the patient demographics for cohorts who received EarlyCDT Lung in papers reporting diagnostic accuracy data.

Table 29 Patient demographics for EarlyCDT Lung studies reporting diagnostic accuracy data

Cohort	Reference	Patient subgroup	Numbers				Age			Gender		Smoker	Assessed cancer risk		Cancer types				Cancer stage	
			Total recruited	Total analysed	N. Cancers	N. No cancer	Mean	Median	Range	N Men	%age male	%age	Mean	Range	N. SCLC	N. NSCLC	N. adeno.	N squamous	Early	Late
Gonzales Retro C-C	Gonzales 2021		NR (Retro C-C)	NR	46	90		63	51.9 - 74.5	96	70	52/48 current v former			3		32	8	37	9
HIPAA	Jett 2014	7-panel	871					61	35 - 95	313	36		4.1% male 1.9 % female							
		All patients																		
HIPAA	Chapman 2012	Clinical population 7-panel	836	836	19	817	60	59	43 - 79 (95%)		36	43.4 curr, 44.3 ex	2.4	0 - 11.9						
HIPAA	Kucera 2012 (CA)	High risk	70	68	15	53														
Lin 2016	Lin 2016 (CA)		31	25	4 (total)	27 (total, to date)	63			14		19 of 31	23%							
Hong Kong	Lau 2017		10	10	5	5	51.5			9		40					5		2	3
EarlyCDT LCS	Jett 2017 (CA)		1235		7		59				45	52 current 48 past			2		2	1		

Table 30 Presents diagnostic accuracy data for cohorts who received EarlyCDT Lung

Table 30 Diagnostic accuracy data reported in EarlyCDT Lung studies

Cohort	Reference	Patient subgroup	Test threshold	N cancers	N no cancer	2x2 data				Sensitivity		Specificity		PPV		NPV	
						TP	FP	TN	FN	Est	95% CI	Est	95% CI	Est	95% CI	Est	95% CI
Gonzales Retro C-C	Gonzales 2021	Susp nodules group	High+Moderate	46	90	6	8	82	40	13	4.9 - 26.3	91.1	83.2 - 96.1				
			High only			6	4	86	40	13	4.9 - 26.3	95.6	89.0 - 98.8				
HIPAA (Jett audit and after)	Massion 2017	Total (7-panel)	Current commerical??	35	131	14	21	108	23					40			
		<4mm				0	5	13	0								
		4 to 20				6	14	73	9					30			
		over 20				8	2	22	14					80			
		30% risk	MAYO model + Aab (both +ve)			10	1	167	30	25		99		91			
			MAYO only			23	25	143	17								
		97% spec	MAYO model + Aab (both +ve)			13	5	163	27								
			MAYO only			5	5	163	35								
	Healy 2017	Nodule set	Commercial	37	111					37.8	22.2 - 53.5	85.6	79.1 - 92.1				
		Nodule set 4-20mm	Commercial	15	87					40	15.2 - 64.8	83.9	76.2 - 91.6				
		Nodule set >20mm	Commercial	22	24					36.4	16.3 - 56.5	91.7	80.6 - 100				

Cohort	Reference	Patient subgroup	Test threshold	N cancers	N no cancer	2x2 data				Sensitivity		Specificity		PPV		NPV	
						TP	FP	TN	FN	Est	95% CI	Est	95% CI	Est	95% CI	Est	95% CI
	Jett 2014	7 panel only	?	35	812	13	70	742	22	37	21 - 55	91	89 - 93	16%			
	Peek 2012	Nodule set		23	68												
HIPAA (Kucera)	Kucera 2012	High risk (some have nodules)		15	53	6	6	47	9	40		89		50%			
HIPAA (Pre nodule subset)	Healey 2013	Cohort "population" set only															
		NLST	two-stratum			5	14	216	13								
			four stratum high/low			2	2	72	4								
			four stratum (middle)			3	12	144	9								
		non NLST	two-stratum			8	40	436	9								
			four stratum high/low			5	4	152	2								
			four stratum (middle)			3	36	284	7								
	Chapman 2012	7-panel clinical population		19	817	9	78	739	10	47		90					
	Healey 2012	*No additonal data*															
Lin 2016	Lin 2016			4	27 (to date)	0	1 est	20 est	4 est	0		95					
Hong Kong	Lau 2017		?	5	5	1	0	5	4								
EarlyCDT LCS	Jett 2017		?	7	345 (pos LDCT only)	2	28	317 (est)	5								
	Phillps 2017			*No further data*													

Cohort	Reference	Patient subgroup	Test threshold	N cancers	N no cancer	2x2 data				Sensitivity		Specificity		PPV		NPV	
						TP	FP	TN	FN	Est	95% CI	Est	95% CI	Est	95% CI	Est	95% CI
ECLS (Scotland)	Sullivan 2021	All patients in EarlyCDT arm		56	6031	18	580	5451	38								

Table 31 presents further diagnostic accuracy data for cohorts who received EarlyCDT Lung

Table 31 Further diagnostic accuracy data reported in EarlyCDT Lung studies

Cohort	Reference	Patient subgroup	Test threshold	AUC		LR+		LR-		Other RR/OR		
				Est	95% CI	Est	95% CI	Est	95% CI	Details	Est	95% CI
Gonzales Retro C-C	Gonzales 2021	Susp nodules group	High+Moderate			1.47	0.54 - 3.98		pos/malig assoc OR	1.54	0.5 - 4.73	
			High only			2.93	0.87 - 9.88			3.22	0.86 - 12.07	
HIPAA (Jett audit and after)	Massion 2017	Total (7-panel)	Current commerical??									
		<4mm										
		4 to 20										
		over 20										
		30% risk	MAYO model + Aab (both +ve)									
			MAYO only									

Cohort	Reference	Patient subgroup	Test threshold	AUC		LR+		LR-		Other RR/OR		
				Est	95% CI	Est	95% CI	Est	95% CI	Details	Est	95% CI
		97% spec	MAYO model + Aab (both +ve)									
			MAYO only									
	Healy 2017	Nodule set	Commercial			2.6	1.4 - 4.8					
		Nodule set 4-20mm	Commercial			2.5	1.1 - 5.4					
		Nodule set >20mm	Commercial			4.4	1.0 - 18.4					
	Jett 2014	7 panel only	*For other data see paper*									
	Peek 2012	Nodule set										
		No additional data										
HIPAA (Kucera)	Kucera 2012											
HIPAA (Pre nodule subset)	Healey 2013		High risk (some have nodules)									
		NLST	Cohort "population" set only									
			two-stratum									
			four stratum high/low									
		non NLST	four stratum (middle)									
	Chapman 2012		two-stratum									

Cohort	Reference	Patient subgroup	Test threshold	AUC		LR+		LR-		Other RR/OR		
				Est	95% CI	Est	95% CI	Est	95% CI	Details	Est	95% CI
	Healey 2012	*No additonal data*	four stratum high/low									
Lin 2016	Lin 2016		four stratum (middle)									
Hong Kong	Lau 2017		7-panel clinical population									
EarlyCDT LCS	Jett 2017											
	Phillps 2017											
ECLS (Scotland)	Sullivan 2021											
			All patients in EarlyCDT arm									
			F*or by stage see paper Table 2*									

11.4 Table of excluded studies with rationale

Table 32 lists studies which were of Early CDT Lung but which did not meet the strict inclusion criteria. Table 33 lists the remaining studies excluded at full-text screening.

Table 32 List of 'near miss' excluded studies

Study	Rationale for exclusion
Boyle (2011)[24]	Exclude based on population
Chapman (2008) [25]	Exclude based on population
Chapman (2010) [161]	Exclude based on population
Chapman (2010) [162]	Exclude based on population
Chapman (2011) [26]	Exclude based on population
Holdenreieder (2011) [163]	Exclude based on population
Lam (2011) [27]	Exclude based on population
MacDonald (2012) [28]	Exclude based on population
MacDonald (2012) [29]	Exclude based on population
McElveen (2016) [164]	Exclude based on population
Murray (2010) [165]	Exclude based on outcome
Peek (2010) [166]	Exclude based on population
Peek (2018) [167]	Exclude based on outcome

Table 33 List of remaining studies excluded at the full-text screening stage

Study	Rationale for exclusion
A Preliminary Study… A Preliminary Study on the Detection of Plasma Markers in Early Diagnosis for Lung Cancer[168]	Exclude based on intervention
Allen (2015) [169]	Exclude based on population
Boyle (2010) [170]	Exclude based on population
Boyle (2010) [171]	Exclude based on population
Boyle (2011)[24]	Exclude based on population
Chang (2019) [172]	Exclude based on intervention
Chapman (2006) [173]	Exclude as could not obtain report
Chapma (2011) [174]	Exclude based on population
Chapman (2008) [25]	Exclude based on population
Chapman (2010) [175]	Exclude based on population
Chapman (2010)[161]	Exclude based on population
Chapman (2010)[162]	Exclude based on population
Chapman (2011)[26]	Exclude based on population
Chapman (2011) [176]	Exclude based on population

Chapman (2012) [7]	Exclude based on population
Chapman (2017) [177]	Exclude based on population
Chest Computed Tomography... Chest Computed Tomography (CT) Screening Study With Antibody Testing[178]	Exclude based on population
Colpitts (2007) [179]	Exclude as could not obtain report
Du (2018) [180]	Exclude based on intervention
EarlyCDT-lung risk assessment... (2012)EarlyCDT-lung risk assessment test (Oncimmune [USA] LLC)	Exclude as could not obtain report
Edelsberg (2018) [106]	Exclude on study design
Eiermann (2011) [181]	Exclude based on population
Farlow (2009) [182]	Exclude based on intervention
Farlow (2010) [183]	Exclude based on intervention
Farlow (2010) [184]	Exclude based on intervention
He (2018) [185]	Exclude based on intervention
Holdenreider (2011) [163]	Exclude based on population
Huang (2020) [186]	
Jamnani (2018) [187]	Exclude based on intervention
Jett (2017) [188]	Exclude based on intervention
Jett (2020) [189]	Exclude based on intervention
Jia (2014) [190]	Exclude based on intervention
Jia (2020) [191]	Exclude based on intervention
Khattar (2010) [192]	Exclude based on intervention
Lam (2011) [27]	Exclude based on population
Lastwika (2018) [193]	Exclude based on intervention
Lastwika (2019) [194]	Exclude based on intervention
Lastwika (2020) [195]	Exclude based on intervention
Lastwika (2020) [196]	Exclude based on intervention
Lu (2019) [197]	Exclude based on intervention
MacDonald (2012) [28]	Exclude based on population
MacDonald (2012) [29]	Exclude based on population
Mathew (2010) [198]	Exclude based on outcome
Mathew (2013) [199]	Exclude based on outcome
Mazzone (2016) [200]	Exclude based on intervention
Mazzone (2018) [201]	Exclude based on outcome
McElveen (2016) [164]	Exclude based on population
Meng (2019) [202]	Exclude based on intervention
Monitoring the Changes… Monitoring the Changes of Tumor-related Biomarkers Before and After Pulmonary Nodule Biopsy[203]	Exclude based on intervention
Mu (2020) [204]	Exclude based on intervention

Murray (2010) [30]	Exclude based on population
Murray (2010) [205]	Exclude based on outcome
Murray (2011) [206]	Exclude based on outcome
Non-Invasive Biomarkers For Early Detection Of Lung Cancers[207]	Exclude based on intervention
Pedchenko (2013) [208]	Exclude based on intervention
Peek (2010) [166]	Exclude based on population
Peek (2010) [209]	Exclude based on outcome
Peek (2018) [167]	Exclude based on outcome
Ren (2015) [210]	Exclude based on intervention
Ren (2015) [211]	Exclude based on intervention
Ren (2018) [212]	Exclude based on intervention
Sutton (2020) [107]	Exclude based on outcome
Trudgen (2014) [213]	Exclude based on intervention
Wang (2020) [214]	Exclude based on intervention
Weycker (2010) [215]	Exclude based on outcome
Weycker (2011) [216]	Exclude based on outcome
Yao (2010) [217]	Exclude based on intervention
Yin-Yu (2019) [218]	Exclude based on intervention
Zhou (2015) [219]	Exclude based on intervention

11.5 Statistical analyses

Figure 25 ROC plot of studies reporting Herder risk

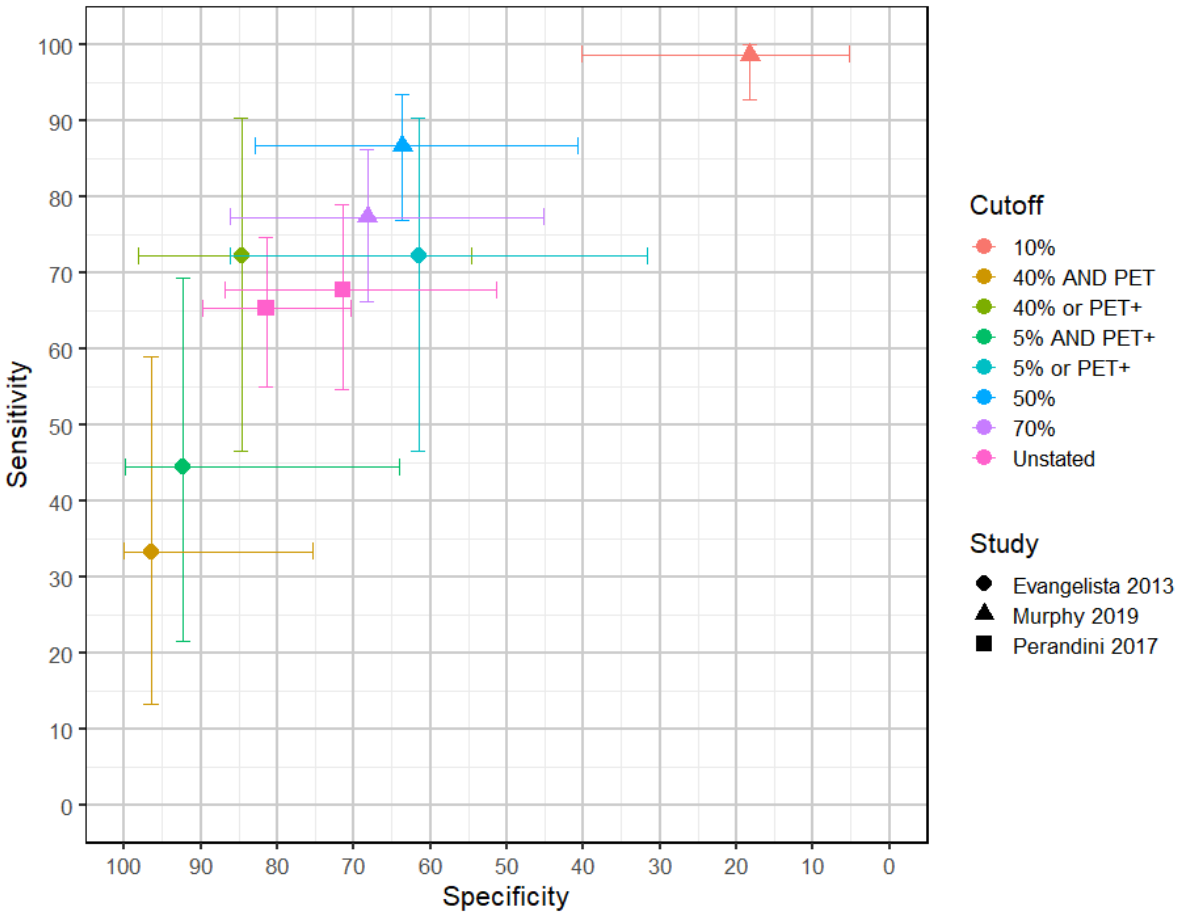


Table 34 Complete results of patient reclassification by risk category in simulation study

Test	Data	Model	Risk group	As proportion of risk group											
				Correctly upgraded			Incorrectly upgraded			Correctly upgraded to >70% risk			Incorrectly upgraded to >70% risk		
				mean	lower 95% CI	upper 95% CI	mean	lower 95% CI	upper 95% CI	mean	lower 95% CI	upper 95% CI	mean	lower 95% CI	upper 95% CI
Brock	Al-Ameri	Healey model	0 to 10%	3.0	1.0	5.2	7.5	3.1	12.4	0.0	0.0	0.0	0.0	0.0	0.0
		EAG model	0 to 10%	2.8	0.0	5.2	5.1	2.1	9.3	0.0	0.0	0.0	0.0	0.0	0.0
Herder	Al-Ameri	Healey model	0 to 10%	1.5	0.0	5.4	3.2	0.0	8.1	0.0	0.0	0.0	0.0	0.0	0.0
			10% to 20%	16.0	0.0	37.5	12.5	0.0	37.5	3.4	0.0	12.5	0.2	0.0	0.0
			20% to 50%	34.9	0.0	71.4	0.3	0.0	0.0	29.0	0.0	57.5	0.3	0.0	0.0
			50% to 70%	27.9	10.0	45.0	6.8	0.0	20.0	27.9	10.0	45.0	6.8	0.0	20.0
		EAG model	0 to 10%	0.0	0.0	0.0	1.2	0.0	5.4	0.0	0.0	0.0	0.0	0.0	0.0
			10% to 20%	16.4	0.0	37.5	12.6	0.0	37.5	0.0	0.0	0.0	0.0	0.0	0.0
			20% to 50%	31.8	0.0	57.1	0.0	0.0	0.0	1.4	0.0	14.3	0.0	0.0	0.0
			50% to 70%	28.6	10.0	45.0	6.9	0.0	20.0	28.6	10.0	45.0	6.9	0.0	20.0
Herder	Perandini	Healey model	0 to 10%	6.2	2.0	11.8	7.3	2.0	13.7	0.0	0.0	0.0	0.0	0.0	0.0
			10% to 20%	20.8	8.6	31.4	10.1	2.9	20.0	3.2	0.0	8.6	0.7	0.0	5.7
			20% to 50%	20.6	5.9	35.3	4.2	0.0	11.8	16.4	0.0	35.3	0.9	0.0	5.9
			50% to 70%	32.3	15.0	50.0	4.9	0.0	15.0	32.3	15.0	50.0	4.9	0.0	15.0
		EAG model	0 to 10%	5.2	0.0	9.8	5.1	0.0	11.8	0.0	0.0	0.0	0.0	0.0	0.0
			10% to 20%	21.1	11.4	31.4	10.2	2.9	20.0	0.0	0.0	0.0	0.0	0.0	0.0
			20% to 50%	8.6	0.0	17.6	2.5	0.0	11.8	0.0	0.0	0.0	0.0	0.0	0.0
			50% to 70%	32.8	15.0	50.0	5.0	0.0	15.0	32.8	15.0	50.0	5.0	0.0	15.0

11.6 Search strategies to identify economic models relevant to lung cancer screening or pulmonary nodules

MEDLINE(R) ALL

via Ovid http://ovidsp.ovid.com/

1946 to March 23, 2021

Searched on 24th March 2021

Retrieved 216 records

Retrieval limited to economic evaluations using a narrow economic search filter developed by CADTH (https://www.cadth.ca/resources/finding-evidence/strings-attached-cadths-database-search-filters#narrow)

1 exp *Lung Neoplasms/ (191984)

2 "Early Detection of Cancer"/ (27568)

3 exp Mass Screening/ (131885)

4 Diagnostic Screening Programs/ (92)

5 2 or 3 or 4 (153125)

6 1 and 5 (4023)

7 ((lung$ or pulmonary) adj3 (neoplas$ or carcinoma$ or cancer$ or tumo?r$) adj4 screen$).ti,ab. (4605)

8 ((NSCLC or SCLC) adj4 screen$).ti,ab. (309)

9 ((earl$ detect$ adj2 cancer$) and (lung$ or pulmonary)).ti,ab. (314)

10 7 or 8 or 9 (5125)

11 6 or 10 (6778)

12 *Solitary Pulmonary Nodule/ (3541)

13 *Multiple Pulmonary Nodules/ (1050)

14 ((lung$ or pulmonary) adj2 (nodule$ or lesion$ or mass or masses)).ti,ab. (26977)

15 ((noncalcified or non calcified) adj2 (nodule$ or lesion$ or mass or masses)).ti,ab. (454)

16 NCPN.ti,ab. (3)

17 (ground-glass adj2 (nodule$ or lesion$ or mass or masses)).ti,ab. (987)

18 (((solid or part-solid or subsolid or sub-solid) adj2 (nodule$ or lesion$ or mass or masses)) and (lung$ or pulmonary)).ti,ab. (1400)

19 ground glass opacit$.ti,ab. (4260)

20 (GGN or GGNs or GGO or GGOs).ti,ab. (1564)

21 (((benign or malignant or indeterminate) adj2 nodule$) and (lung$ or pulmonary)).ti,ab. (1381)

22 coin lesion$.ti,ab. (484)

23 (IPN or IPNs).ti,ab. (1735)

24 or/12-23 (35043)

25 11 or 24 (40796)

26 *economics/ (10730)

27 exp *"costs and cost analysis"/ (73688)

28 (cost minimi* or cost-utilit* or health utilit* or economic evaluation* or economic review* or cost outcome or cost analys?s or economic analys?s or budget* impact analys?s).ti,ab,kf,kw. (34621)

29 (cost-effective* or pharmacoeconomic* or pharmaco-economic* or cost-benefit or costs).ti,kf,kw. (75866)

30 (life year or life years or qaly* or cost-benefit analys?s or cost-effectiveness analys?s).ab,kf,kw. (32132)

31 (cost or economic*).ti,kf,kw. and (costs or cost-effectiveness or markov).ab. (60556)

32 (economic adj2 model*).mp. (13611)

33 26 or 27 or 28 or 29 or 30 or 31 or 32 (186355)

34 25 and 33 (281)

35 (editorial or historical article or letter).pt. (2032515)

36 34 not 35 (263)

37 limit 36 to english language (241)

38 limit 37 to yr="2000 -Current" (222)

39 exp animals/ not humans.sh. (4804106)

40 38 not 39 (216)

Key:

/ = subject heading (MeSH heading)

sh = subject heading (MeSH heading)

exp = exploded subject heading (MeSH heading)

* = focus applied to subject heading - retrieves only those articles where subject heading is primary focus of the article

$ = truncation

? = optional wildcard – stands for one or no characters

ti,ab = terms in title or abstract fields

kf = author keyword field

mp = multi-purpose field – includes searching of title, abstract, subject headings, other title, author keywords, synonyms

adj3 = terms within three words of each other (any order)

pt = publication type

Embase

via Ovid http://ovidsp.ovid.com/

1974 to 2021 March 24

Searched on 24th March 2021

Retrieved 539 records

Retrieval limited to economic evaluations using a narrow economic search filter developed by CADTH (https://www.cadth.ca/resources/finding-evidence/strings-attached-cadths-database-search-filters#narrow)

1 exp *lung tumor/ (222641)

2 early cancer diagnosis/ (8031)

3 *mass screening/ (22493)

4 *cancer screening/ (32117)

5 *screening/ (33181)

6 2 or 3 or 4 or 5 (94089)

7 1 and 6 (4302)

8 ((lung$ or pulmonary) adj3 (neoplas$ or carcinoma$ or cancer$ or tumo?r$) adj4 screen$).ti,ab. (7350)

9 ((NSCLC or SCLC) adj4 screen$).ti,ab. (673)

10 ((earl$ detect$ adj2 cancer$) and (lung$ or pulmonary)).ti,ab. (496)

11 8 or 9 or 10 (8333)

12 7 or 11 (9289)

13 *lung nodule/ (6861)

14 lung coin lesion/ (560)

15 *multiple pulmonary nodules/ (248)

16 ((lung$ or pulmonary) adj2 (nodule$ or lesion$ or mass or masses)).ti,ab. (41212)

17 ((noncalcified or non calcified) adj2 (nodule$ or lesion$ or mass or masses)).ti,ab. (717)

18 NCPN.ti,ab. (4)

19 (ground-glass adj2 (nodule$ or lesion$ or mass or masses)).ti,ab. (1589)

20 (((solid or part-solid or subsolid or sub-solid) adj2 (nodule$ or lesion$ or mass or masses)) and (lung$ or pulmonary)).ti,ab. (2539)

21 ground glass opacit$.ti,ab. (7787)

22 (GGN or GGNs or GGO or GGOs).ti,ab. (2524)

23 (((benign or malignant or indeterminate) adj2 nodule$) and (lung$ or pulmonary)).ti,ab. (2366)

24 coin lesion$.ti,ab. (463)

25 (IPN or IPNs).ti,ab. (2283)

26 or/13-25 (54210)

27 12 or 26 (61910)

28 *economics/ (26847)

29 economic evaluation/ or "cost benefit analysis"/ or "cost effectiveness analysis"/ or "cost minimization analysis"/ or "cost utility analysis"/ (249213)

30 (cost minimi* or cost-utilit* or health utilit* or economic evaluation* or economic review* or cost outcome or cost analys?s or economic analys?s or budget* impact analys?s).ti,ab,kw. (54310)

31 (cost-effective* or pharmacoeconomic* or pharmaco-economic* or cost-benefit or costs).ti,kw. (113460)

32 (life year or life years or qaly* or cost-benefit analys?s or cost-effectiveness analys?s).ab,kw. (50521)

33 (cost or economic*).ti,kw. and (costs or cost-effectiveness or markov).ab. (96734)

34 (economic adj2 model*).mp. (8150)

35 28 or 29 or 30 or 31 or 32 or 33 or 34 (370615)

36 27 and 35 (976)

37 (conference abstract or "conference review").pt. (4079881)

38 36 not 37 (756)

39 (editorial or letter).pt. (1862035)

40 38 not 39 (651)

41 limit 40 to english language (592)

42 limit 41 to yr="2000 -Current" (545)

43 (animal/ or animal experiment/ or animal model/ or animal tissue/) not exp human/ (3889280)

44 42 not 43 (539)

Key:

/ = subject heading (Emtree heading)

sh = subject heading (Emtree heading)

exp = exploded subject heading (Emtree heading)

* = focus applied to subject heading - retrieves only those articles where subject heading is primary focus of the article

$ = truncation

? = optional wildcard – stands for one or no characters

ti,ab = terms in title or abstract fields

kw = author keyword field

mp = multi-purpose field – includes searching of title, abstract, subject headings, other title, keywords, synonyms

adj3 = terms within three words of each other (any order)

pt = publication type

NHS Economic Evaluations Database (NHS EED)

Health Technology Assessment (HTA) database

via http://www.crd.york.ac.uk/CRDWeb/

Searched on 24th March 2021

Retrieved NHS EED 59 records and HTA 28 records (pre-2000 records removed in EndNote)

1 (MeSH DESCRIPTOR Lung Neoplasms EXPLODE ALL TREES) 1151

2 ((lung* or pulmonary) adj3 (neoplas* or carcinoma* or cancer* or tumor* or tumour*)) 1428

3 ((neoplas* or carcinoma* or cancer* or tumor* or tumour*) adj3 (lung* or pulmonary)) 856

4 (NSCLC or SCLC) 284

5 #1 OR #2 OR #3 OR #4 1473

6 MeSH DESCRIPTOR Diagnostic Screening Programs EXPLODE ALL TREES 0

7 MeSH DESCRIPTOR Mass Screening EXPLODE ALL TREES 2347

8 MeSH DESCRIPTOR Early Detection of Cancer 277

9 (screen*) 8160

10 ((earl* detect* adj2 cancer*) OR (cancer* adj2 earl* detect*)) 358

11 #6 OR #7 OR #8 OR #9 OR #108254

12 #5 AND #11 176

13 (MeSH DESCRIPTOR Solitary Pulmonary Nodule) 27

14 (MeSH DESCRIPTOR Multiple Pulmonary Nodules) 1

15 ((lung* or pulmonary) adj2 (nodule* or lesion* or mass or masses)) OR ((nodule* or lesion* or mass or masses) adj2 (lung* or pulmonary)) 117

16 ((noncalcified or non-calcified) adj2 (nodule* or lesion* or mass or masses)) OR ((nodule* or lesion* or mass or masses) adj2 (noncalcified or non-calcified)) 6

17 (NCPN) 0

18 (ground-glass adj2 (nodule* or lesion* or mass or masses)) 0

19 ((nodule* or lesion* or mass or masses) adj2 ground-glass) 0

20 ((solid or part-solid or subsolid or sub-solid) adj2 (nodule* or lesion* or mass or masses))24

21 ((nodule* or lesion* or mass or masses)) adj2 ((solid or part-solid or subsolid or sub-solid)) 1

22 (lung* or pulmonary) 4973

23 #20 OR #21 24

24 #22 AND #23 1

25 ("ground glass opacity" OR "ground glass opacities") 2

26 (GGN or GGNs or GGO or GGOs) 0

27 ((benign or malignant or indeterminate) adj2 nodule*) OR (nodule* adj2 (benign or malignant or indeterminate)) 53

28 #27 AND #22 19

29 ("coin lesion") OR ("coin lesions") 1

30 (IPN or IPNs) 0

31 #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #24 OR #25 OR #26 OR #28 OR #29 OR #30 123

32 #12 OR #31 260

33 (*) IN NHSEED 17613

34 #32 AND #33 76

35 (*) IN HTA 17351

36 #32 AND #35 58

37 MeSH DESCRIPTOR Economics 23

38 MeSH DESCRIPTOR Costs and Cost Analysis EXPLODE ALL TREES 17164

39 (cost* or economic* or pharmacoeconomic* or pharmaco-economic* or markov or budget* or "life year" or "life years" or qaly*) 26569

40 #37 OR #38 OR #39 26573

41 #36 AND #40 28 (HTA results)

42 (#32 AND #33) FROM 2000 TO 2021 59 (NHS EED results)

Key:

MeSH DESCRIPTOR = subject heading (MeSH heading)

* = truncation

adj3 = terms within three words of each other (order specified)

International Health Technology Assessment (INAHTA) database

https://database.inahta.org/

Searched on 25th March 2021

Retrieved 38 records

(((cost* or economic* or pharmacoeconomic* or "pharmaco-economic" or "pharmaco-economics" or markov or budget* or "life year" or "life years" or qaly*)[Title] OR (cost* or economic* or pharmacoeconomic* or "pharmaco-economic" or "pharmaco-economics" or markov or budget* or "life year" or "life years" or qaly*)[abs] OR (cost* or economic* or pharmacoeconomic* or "pharmaco-economic" or "pharmaco-economics" or markov or budget* or "life year" or "life years" or qaly*)[Keywords]) OR ("Costs and Cost Analysis"[mhe]) OR ("Economics"[mh])) AND ((((("coin lesion" or "coin lesions" or IPN or IPNs)[Title] OR ("coin lesion" or "coin lesions" or IPN or IPNs)[abs] OR ("coin lesion" or "coin lesions" or IPN or IPNs)[Keywords]) OR (((nodule*)[Title] OR (nodule*)[abs] OR (nodule*)[Keywords]) AND ((benign or malignant or indeterminate)[Title] OR (benign or malignant or indeterminate)[abs] OR (benign or malignant or indeterminate)[Keywords])) OR (("ground glass opacity" or "ground glass opacities" or GGN or GGNs or GGO or GGOs)[Title] OR ("ground glass opacity" or "ground glass opacities" or GGN or GGNs or GGO or GGOs)[abs] OR ("ground glass opacity" or "ground glass opacities" or GGN or GGNs or GGO or GGOs)[Keywords]) OR (((solid or "part-solid" or "part solid" or subsolid or "sub-solid" or "sub solid")[Title] OR (solid or "part-solid" or "part solid" or subsolid or "sub-solid" or "sub solid")[abs] OR (solid or "part-solid" or "part solid" or subsolid or "sub-solid" or "sub solid")[Keywords]) AND ((nodule* or lesion* or mass or masses)[Title] OR (nodule* or lesion* or mass or masses)[abs] OR (nodule* or lesion* or mass or masses)[Keywords]) AND ((lung* or pulmonary)[Title] OR (lung* or pulmonary)[abs] OR (lung* or pulmonary)[Keywords])) OR ((NCPN)[Title] OR (NCPN)[abs] OR (NCPN)[Keywords]) OR

(((noncalcified or "non-calcified" or "non calcified" or "ground-glass" or "ground glass")[Title] OR (noncalcified or "non-calcified" or "non calcified" or "ground-glass" or "ground glass")[abs] OR (noncalcified or "non-calcified" or "non calcified" or "ground-glass" or "ground glass")[Keywords]) AND ((nodule* or lesion* or mass or masses)[Title] OR (nodule* or lesion* or mass or masses)[abs] OR (nodule* or lesion* or mass or masses)[Keywords])) OR (((nodule* or lesion* or mass or masses)[Title] OR (nodule* or lesion* or mass or masses)[abs] OR (nodule* or lesion* or mass or masses)[Keywords]) AND ((lung* or pulmonary)[Title] OR (lung* or pulmonary)[abs] OR (lung* or pulmonary)[Keywords])) OR ("Multiple Pulmonary Nodules"[mh]) OR ("Solitary Pulmonary Nodule"[mh])) OR (((((cancer*)[Title] OR (cancer*)[abs] OR (cancer*)[Keywords]) AND ((earl* detect*)[Title] OR (earl* detect*)[abs] OR (earl* detect*)[Keywords])) OR ((screen*)[Title] OR (screen*)[abs] OR (screen*)[Keywords]) OR ("Early Detection of Cancer"[mh]) OR ("Mass Screening"[mhe]) OR ("Diagnostic Screening Programs"[mh])) AND (((NSCLC or SCLC)[Title] OR (NSCLC or SCLC)[abs] OR (NSCLC or SCLC)[Keywords]) OR (((neoplas* or carcinoma* or cancer* or tumor* or tumour*)[Title] OR (neoplas* or carcinoma* or cancer* or tumor* or tumour*)[abs] OR (neoplas* or carcinoma* or cancer* or tumor* or tumour*)[Keywords]) AND ((lung* or pulmonary)[Title] OR (lung* or pulmonary)[abs] OR (lung* or pulmonary)[Keywords])) OR ("Lung Neoplasms"[mhe])))

date limit applied: 2000-2021

Key:

[Keywords] = search of keywords field

[abs] = search of abstract field

[Title] = search of title field

[mh] = subject heading search

[mhe] = exploded subject heading search

* = truncation

EconLit

via Ovid http://ovidsp.ovid.com/

1886 to March 18, 2021

Searched on 25th March 2021

Retrieved 5 records

1 ((lung$ or pulmonary) adj3 (neoplas$ or carcinoma$ or cancer$ or tumo?r$) adj4 screen$).mp. (4)

2 ((NSCLC or SCLC) adj4 screen$).mp. (0)

3 ((earl$ detect$ adj2 cancer$) and (lung$ or pulmonary)).mp. (1)

4 ((lung$ or pulmonary) adj2 (nodule$ or lesion$ or mass or masses)).mp. (0)

5 ((noncalcified or non calcified) adj2 (nodule$ or lesion$ or mass or masses)).mp. (0)

6 NCPN.mp. (0)

7 (ground-glass adj2 (nodule$ or lesion$ or mass or masses)).mp. (0)

8 (((solid or part-solid or subsolid or sub-solid) adj2 (nodule$ or lesion$ or mass or masses)) and (lung$ or pulmonary)).mp. (0)

9 ground glass opacit$.mp. (0)

10 (((benign or malignant or indeterminate) adj2 nodule$) and (lung$ or pulmonary)).mp. (0)

11 coin lesion$.mp. (0)

12 or/1-11 (5)

Key:

$ = truncation

? = optional wildcard – stands for one or no characters

mp = multi-purpose field – includes searching of title, abstract, subject headings, other title, keywords, synonyms

adj3 = terms within three words of each other (any order)

11.7 Critical appraisal of cost-effectiveness studies of EarlyCDT Lung

Table 35 Yang et al. Checklist for model-based economic evaluations of diagnostic tests

	Response (Y,N or NA)	Comments
1. Decision problem and scope specified		
1. Is there a clear statement of the decision problem?	Y	
2. Is the perspective of the model stated clearly?	Y	
3. Has the target population been identified?	Y	
4. Are the model inputs consistent with the stated perspective?	Y	
5. Are the primary outcomes of the model consistent with the perspective, scope and overall objective of the model?	Y	
2. Identification and description of the comparators		
6. Have all the feasible and practical options been identified?	N	It is not discussed whether there were other feasible and relevant alternatives
7. Have the comparators being evaluated been clearly described?	N	It is unclear how patients are managed following identification
8. If comparators have been excluded from the evaluation, have these exclusions been justified?	NA	
3. Appropriate data identification		
9. Are the data identification methods transparent, systematic and appropriate given the objectives of the model?	N	
4. Sufficient detail for data incorporation		

10. Have all data incorporated into the model been described and referenced in sufficient detail?	N	There is not sufficient detail to understand which data was extracted from each of the sources referenced to parameterise life-expectancy projections.
11. Where choices have been made between data sources, are these justified appropriately?	N	
12. Are transition probabilities calculated appropriately?	NA	Not enough detail to assess this
13. Has discounting been conducted?	Y	
5. Quality and incorporation of test accuracy data		
14. Has the quality of the test accuracy data been assessed?	N	
15. Have diagnostic accuracy data been derived from high quality data sources (hierarchy of evidence)?		Single source of data to inform data accuracy is not described in sufficient detail to establish quality of data
16. Are tests in sequence treated dependently, where appropriate?	N	No comment on dependency between tests in a diagnostic sequence
6. Quality and incorporation of treatment data		
17. Has the quality of the treatment effect data been assessed?	N	
18. Have relative treatment effects been derived from high quality data sources (hierarchy of evidence)?	N	Treatment does not seem to have been explicitly modelled. Text suggests that life expectancy is conditional on disease stage rather than treatment.
7. Source and incorporation of cost data		
19. Has the source of cost data been presented clearly?	Y	
20. Have costs been inflated to a specific year, where appropriate?	Y	
8. Source and incorporation of utility data		
21. Is the source for the utility weights referenced and justified?	Partly	Referenced, but not justified how these were identified and selected
22. Are the utilities incorporated into the model appropriately?	N	Not sufficient detail in the paper to assess this properly, but it seems that only utilities for patients without malignancy were age (and gender) dependents.

9. Model structure		
23. Have the reasons behind the type of decision analytic model chosen been fully described and justified?	N	
24. Has a systematic review of existing economic evaluations been carried out?	N	
25. Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	NA	The structure of the model is not sufficiently described or depicted to assess whether it is consistent with the health condition.
26. Are the structural assumptions underpinning the model transparent and justified?	N	Model structure not described
27. Have the methods used to extrapolate short-term results to final outcomes been documented and justified?	N	It is unclear how and at what point in the model the long-term extrapolation was done
28. Has the time horizon been stated and justified?	N	The choice of outcomes suggests that it is a life time model, but this is not clearly stated
29. Has cycle length of Markov models been justified?	N	Probabilities are described as monthly, which suggests a cycle length of one month. No justification provided.
10. Uncertainty		
30. Has parameter uncertainty been addressed via sensitivity analysis?	Y	
31. Has probabilistic sensitivity analysis been carried out? If not, has this omission been justified?	N	No justification for not conducting probabilistic sensitivity analysis
32. If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	N	Only states that reasonable alternative values were used
33. If data have been incorporated as distributions, has the choice of distribution for each parameter been described and justified?	NA	
34. Have structural uncertainties been addressed via sensitivity analysis?	Partly	Alternative CT surveillance schedule is the only structural assumption tested

35. Have alternative assumptions related to final outcomes been explored through sensitivity analysis?	N	
36. Has value of information analysis been done?	N	
11. Validity		
37. Has the face validity been reviewed by someone external to the model developers?	N	Not described
38. Has the mathematical logic of the model been assessed? (e.g. using null and extreme values)	N	Not described
39. Have the model and its results been compared to the findings of other models and studies, and any disagreements or inconsistencies been explained (cross-validity)?	Partly	Volume doubling time and risk of progression were compared to external data and found consistent

Table 36 Yang et al. Checklist for model-based economic evaluations of diagnostic tests: Sutton et al, 2020

	Response (Y,N or NA)	Comments
1. Decision problem and scope specified		
40. Is there a clear statement of the decision problem?	Y	
41. Is the perspective of the model stated clearly?	Y	
42. Has the target population been identified?	Partly	But it is unclear why it was considered the relevant population
43. Are the model inputs consistent with the stated perspective?	Y	
44. Are the primary outcomes of the model consistent with the perspective, scope and overall objective of the model?	Y	
2. Identification and description of the comparators		
45. Have all the feasible and practical options been identified?	N	No discussion of relevant comparators
46. Have the comparators being evaluated been clearly described?	Y	
47. If comparators have been excluded from the evaluation, have these exclusions been justified?	NA	
3. Appropriate data identification		
48. Are the data identification methods transparent, systematic and appropriate given the objectives of the model?	N	The authors state that "*Rather than doing an extensive systematic review to identify the best available evidence to populate the model, this study has made extensive use of the parameters, data and model structure from the study by Gould et al., 2003*". It is unclear why this was considered appropriate.
4. Sufficient detail for data incorporation		
49. Have all data incorporated into the model been described and referenced in sufficient detail?	Y	

50. Where choices have been made between data sources, are these justified appropriately?	N	This is not discussed
51. Are transition probabilities calculated appropriately?	NA	Not enough detail to assess this
52. Has discounting been conducted?	Y	
5. Quality and incorporation of test accuracy data		
53. Has the quality of the test accuracy data been assessed?	N	
54. Have diagnostic accuracy data been derived from high quality data sources (hierarchy of evidence)?	?	Single source of data to inform data accuracy is not described in sufficient detail to establish quality of data
55. Are tests in sequence treated dependently, where appropriate?	N	No comment on dependency between tests in a diagnostic sequence
6. Quality and incorporation of treatment data		
56. Has the quality of the treatment effect data been assessed?	N	
57. Have relative treatment effects been derived from high quality data sources (hierarchy of evidence)?	N	Treatment effects are not applied as relative effects. Patient outcomes are conditional on disease stage at which patients are diagnosed.
7. Source and incorporation of cost data		
58. Has the source of cost data been presented clearly?	Y	
59. Have costs been inflated to a specific year, where appropriate?	Y	
8. Source and incorporation of utility data		
60. Is the source for the utility weights referenced and justified?	Partly	Referenced, but not justified how these were identified and selected. Some utilities are taken from a study on detection of liver fibrosis
61. Are the utilities incorporated into the model appropriately?	?	Not sufficient detail in the paper to assess this
9. Model structure		

62. Have the reasons behind the type of decision analytic model chosen been fully described and justified?	Partly	Only for the Markov model component.
63. Has a systematic review of existing economic evaluations been carried out?	N	
64. Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	Y	
65. Are the structural assumptions underpinning the model transparent and justified?	Partly	Mostly, but unclear how the diagnostic accuracy of CT scans was implemented in the model
66. Have the methods used to extrapolate short-term results to final outcomes been documented and justified?	N	
67. Has the time horizon been stated and justified?	Y	The choice of outcomes suggests that it is a life time model, but this is not clearly stated
68. Has cycle length of Markov models been justified?	N	
10. Uncertainty		
69. Has parameter uncertainty been addressed via sensitivity analysis?	N	Only diagnostic accuracy of EarlyCDT and cost of the this test were varied in sensitivity analysis
70. Has probabilistic sensitivity analysis been carried out? If not, has this omission been justified?	Y	
71. If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	N	Only states that reasonable alternative values were used
72. If data have been incorporated as distributions, has the choice of distribution for each parameter been described and justified?	N	
73. Have structural uncertainties been addressed via sensitivity analysis?	N	

74. Have alternative assumptions related to final outcomes been explored through sensitivity analysis?	N	
75. Has value of information analysis been done?	Y	EVPI and EVPPI. Unclear how the parameters considered in the EVPPI analysis were aggregated
11. Validity		
76. Has the face validity been reviewed by someone external to the model developers?	N	Not described
77. Has the mathematical logic of the model been assessed? (e.g. using null and extreme values)	N	Not described
78. Have the model and its results been compared to the findings of other models and studies, and any disagreements or inconsistencies been explained (cross-validity)?	N	

11.8 Additional reviews to support model conceptualisation

11.8.1 Review of cost-effectiveness studies on other diagnostics for lung cancer diagnosis

In the review of cost-effectiveness studies on other diagnostics for lung cancer, we did not consider in much detail the diagnostic strategies implemented and their accuracy (which are context specific). Instead, we focussed on identifying the assumptions and evidence supporting quantification of the value components that could be of relevance for a future assessment of EarlyCDT Lung, namely:

- was increased detection of lung cancer in relation to surveillance considered?
- was early diagnosis the key mechanism of value?,
- was overdiagnosis/overtreatment considered?, and
- how were false positives assumed to be managed?

We also considered the assumptions and evidence supporting linkage to long terms health and costs outcomes of these, namely:

- how was earlier diagnosis linked to progression of disease – stage shift,
- how was stage shift linked to improved long term outcomes of treatment – outcomes component.

11.8.1.1 Overview of the diagnostic models

From the identified studies, we extracted the assumptions and evidence supporting quantifications of the value components related to misclassification introduced by the tests in the diagnostic pathways. Table 37 identifies the studies where these features were quantified.

Table 37 Diagnostic studies summary: Identification of value components related to classification

Study (year)	Surveillance strategy modelled?	Earlier diagnosis?	Increased detection?	Overtreatment of indolent malignant or decision to treat benign?	False positives allowed?	Other
D'Andrea (2020)	Yes	Yes	Yes	No	Yes	No
Deppen (2014)	Yes	Yes	No	Yes	No	No
Dietlein (2000)	Yes	Yes	Yes	Yes	No	No
Goehler (2014)	Yes	Yes	Yes	NR	NR	Regression of benign nodules
Gould (2003)	Yes	Yes	No	No	Yes	No
Jiang (2020)	Yes	Yes	Yes	No	NR	No
Lejeune (2005)	Yes	Yes	Yes	No	Yes	Regression of benign nodules

Rickets (2020)	No	Yes	Yes	No	Yes	No

*assumptions/evidence on CT surveillance accuracy not reported

NR, not reported

All diagnostic studies ascribed value from earlier SPN diagnosis (in common with the EarlyCDT Lung studies) but, additionally, i) two studies[113, 114] considered a decision to treat of benign nodules explicitly (but not the overtreatment of indolent malignant nodules), ii) two studies [113, 114]did not allow for false positives at the end of the diagnostic strategy (i.e., surgical treatment of false positives was not allowed) and iii) two studies [115, 117] allowed for regression of benign nodules with potential early discharge from surveillance. Specifically how these were considered in the studies is described below in further detail.

11.8.1.2 Evidence linkage for earlier/higher detection of lung cancer

Table 38 summarises in further detail the evidence linkage regarding earlier and/or higher detection of lung cancer.

Studies considered the possibility of higher detection variably. One study[109] assumed 100% sensitivity for surveillance and assumed no missed cancers across all diagnostic strategies analysed. Two studies[115, 117] explicitly considered an imperfect sensitivity for surveillance, raising the possibility of higher detection for diagnostic strategies that reduce the proportion of individuals undergoing surveillance (Goehler et al., 2014 conditioned the sensitivity of surveillance on nodule size, location, whether first CT or follow-up). Two studies[112, 114]assumed that some malignant nodules remain undetected under CT surveillance, but it is unclear how this was parameterised (with the exception of the proportion of patients who do not uptake CT surveillance in Dietlien et al., 2000[114]); there seems to be an implicit assumption that that the specificity of CT surveillance is lower than 100%. One study[116] considered higher detection for one diagnostic strategy vs. the alternative, rather than against CT surveillance. This was because the former allowed incidental identification of nodules anywhere lung, whereas the latter could only identify nodules in the lower and mid fields. The authors did not, however, provide sufficient detail to characterise the diagnostic accuracy of follow-up tests (invasive and non-invasive) after incidental identification of the nodules.

Table 38 Modelling of value components relating to earlier/increased detection of lung cancer - diagnostic studies summary

Study (year)	Increased detection		Delay to diagnosis			Disease stage at diagnosis, conditional on delay	
	False negatives allowed for surveillance?	source	Intervention causing delay	Mechanism	Source	Mechanism	Source
D'Andrea (2020)	Some cancers remain undetected	NR	CTSurv	probabilities of detection at different timepoints (unclear)	NR	Assumption: All stage I progress to stage II	NR
Deppen (2014)	No (text suggests 100% specificity for CT surv	NR	CTSurv	single time point for delay to diagnosis.	Assumption based on Gambhir, 1998[220]	Unclear	NR
Dietlein (2000)	1. Some cancers remain undetected* and 2. uptake of CTsurv is imperfect	1. Seely, 1993 [221] 2. NR	CTSurv	probabilities of detection at different timepoints	Assumption, based on assumed mean VDT[220, 222]	Unclear	NR
Goehler (2014)	Imperfect sensitivity of CT surve	Swensen, 2003[223]	no follow-up, CTSurv	probabilities of detection at different timepoints	Model: Natural history model simulating growth and progression[154]		
Gould (2003)	No (100% sensitivity of serial chest radiographs)	Assumption, no justification	CTSurv	probabilities of detection at different timepoints	Model: of distribution of VDT based on Steele et al., 1973[110]	Preclinical progression model	Unclear, informed by VDT data[110]
Jiang (2020)	NR		CTCS	Unclear	Unclear	Assumption: (unclear)	NR
Lejeune (2005)	Yes (imperfect sensitivity of CTsurv)	Zwirewich, 1991[224], Swensen, 1996[225]	CTSurv	probabilities of detection at different timepoints	Literature[226, 227] (based on VDT)**	Assumption: All stage 1 progress to stage 2	NR
Rickets (2020)	NA	NA	FNs to diagnostic tests	single time point for delay to diagnosis.	Assumption, justification NR	Preclinical progression model	Expert opinion[155]

CTCS, conventional computed tomographic calcium scoring; CT-FNA, CT guided fine-needle aspiration; CTsurv, CT surveillance; FN, false negatives; NA, not applicable; NR, not reported.

*the authors state that some stage N2/N3 cancers remain undetected, but may only apply to PET-CT; ** values used could not be found in source references

All studies modelled a delay to diagnosis (mostly for the strategies with an element of CT surveillance or equivalent serial imaging), but the way in which this was implemented varied across studies. In two studies, [112, 116] it was unclear how the delay to diagnosis was implemented in the model. Two other studies defined the delay to diagnosis by assuming that late diagnosis occurred at a single point in time of one month (6 months in a scenario analysis) [113]or two months.[118] None of these studies justify their assumption. In the remaining studies, diagnosis occurred across multiple points in time, with the probability of detection at each time point informed either by assumptions, [114, 117] or by explicit modelling of nodule growth.[109, 115]

Dietlien et al., 2000, [114] assumed that 50% of patients with malignant nodules would be detected after three months of CT surveillance and the rest after 6 months; this was based on an assumed mean VDT of three months and an implicit assumption that CT surveillance has a 100% sensitivity to detect nodule growth. In this study, diagnostic accuracy of CT surveillance to detect mediastinal involvement is also assumed to be imperfect with a percentage of N2/N3 cancers (TNM stage classification, where indicates N refers to the number of nearby lymph nodes that have cancer) going undetected. These assumptions were not justified.

Lejeune et al., 2005, [117] assumed a cumulative malignant nodule growth rate of 50% during the first three months, 75% at six months, 90% at nine months, and 100% at one year, which were sourced from a 1986 study[227] These rates were combined with CT surveillance diagnostic accuracy estimates to determine the probability of growth being detected by CT surveillance each three-month cycle.

Goehler et al., 2014, [115] used a microsimulation model, the Lung Cancer Policy Model (LCPM)[154]to simulate nodule growth according to patient characteristics, flow across the diagnostic pathway and subsequent management. Only the structure of the diagnostic component of the simulation model is described in the paper, so it is not clear how the nodule growth is modelled. Diagnostic accuracy of CT scans was conditional on nodule size and location (central vs. peripheral), and type of CT (initial CT scan which incidentally identifies nodule vs. routine CT scan as part of CT surveillance). The diagnostic accuracy of CT to detect growth combined with the simulated nodule growth over time determined the delay to diagnosis of malignant nodules.

In Gould et al., 2003,[109] malignant nodules were modelled to double in size every 5.24 months, the mean VDT for a distribution of observed doubling times for 67 pulmonary nodules and mass lesions (measured based on chest radiographs) from the Veterans Administration–Armed Forces Cooperative Study on Asymptomatic Pulmonary Nodules, [110] which was also used to inform he EarlyCDT Lung cost-effectiveness studies.[106, 107] It was assumed that chest radiographs in the (watchful waiting component of the diagnostic strategies) were 100% sensitive to detect tumour growth, defined as one doubling in tumour volume or a change in nodule size from 2 cm to 2.5 cm in diameter. The

diagnostic accuracy of watchful waiting was used to inform transitions in a Markov model between undiagnosed and diagnosed health states for patients who did not have a correct diagnose at the end of a decision tree used to characterise the diagnostic pathway. Although the text suggests that VDT was formally modelled, the way in which this was implemented in the model is not completely clear.

The delay to diagnosis was linked to cancer stage at diagnosis in four studies.[109, 112, 117, 118] Rickets et al., 2020, [118] modelled disease progression for undiagnosed patients over time, as a set of sequential health states (Stage I to Stage IV). Transition probabilities between stages were informed by elicited evidence from a study on early lung diagnosis cancer promoted by public health policies on disease awareness.[155] In Gould et al., 2003, [109] the disease progression across stages (local, regional and distant cancer) is assumed to depend on VDT. It is unclear how the probabilities of disease progression were derived, but the text suggests that calibration was used over the VDT data in Steele et al., 1973, [110] and an assumption of equal probability of progression for local to regional disease and for regional to distant disease. The linkage mechanism between diagnostic delay and disease stage progression in D'Andrea et al., 2020, [112] and in Lejeune et al., 2005,[117] is not explicit but appears to rely on assumptions. In both papers, undetected malignancies are assumed to progress from stage I to stage II at the end of the first year [112] or for the duration of surveillance[117], without justifying such assumption. The link between delay to diagnosis and disease progression is even less well characterised in the remaining diagnostics studies. In two studies[113, 114] survival is modelled conditional on disease stages, which suggests an assumed link between delay and disease progression. Ghoeler et al., 2014, [115] state that the microsimulation model captures disease progression alongside nodule growth, but the model is not described in the manuscript. One study[116] did not present sufficient information to understand what the timing of the delay was and the stage distribution for identified vs. unidentified malignant nodules.

Table 39 Modelling of link between disease status and staging, and outcomes

Study (year)	Model for outcomes?	Staging categorisation for malignant	Long-term outcomes	Disease status	
				Malignant	Benign
				Conditional on:	
D'Andrea (2020)	Yes	Stage 1,2	Survival	Staging, age	--
			HRQoL	Staging	Age
			Costs	Treatment (surgical)	--
Deppen (2014)	No, LE payoff	Stage 1, 2, 3/4	Survival	Staging, age	Age
			HRQoL	Staging (unclear)	--
			Costs	Treatment (surgical)	--
Dietlein (2000)	No, LE payoff	Stage T1N0, T1 or T2N0/1, T(any) N2/3	Survival	Staging	--
			HRQoL	Not modelled	Not modelled
			Costs	Treatment (surgical, palliative)	--

Goehler (2014)	Yes	NR	Survival	Staging (unclear), comorbidity (CAD)	NR
			HRQoL	Staging, histology, type of treatment and response, recurrence, time post-treatment	Age, sex
			Costs	NR	NR
Gould (2003)	Yes	Local, regional, distant	Survival	Staging, age	Age
			HRQoL	Staging, recurrence, age, sex	Age, sex
			Costs	Treatment (surgical)	NA
Jiang (2020)	NR	NR	Survival	Staging (unclear)	--
			HRQoL	Not modelled	Not modelled
			Costs	Cancer treatment	NA
Lejeune (2005)	LE payoff	T1, T2	Survival	Staging (NR), age	Age
			HRQoL	Not modelled	Not modelled
			Costs	NR	NR
Rickets (2020)	Model	Stage 1-4	Survival	Staging, age	Age, sex
			HRQoL	Staging, age	Age
			Costs	Staging, delayed diagnosis	NR

CAD, coronary artery disease; HRQoL, health-related quality of life; LE, life expectancy; NA, not applicable; NR, not reported

All studies appear to condition the survival outcomes of patients with lung cancer on staging (although not all explicitly state it [e.g., Ghoeler et al., 2014[115]]), consistently with the use of the mechanism of linking early (and increased) diagnosis to stage shift. Survival for these patients is also often conditioned on age. One study[115] incorporated a competing mortality risks due to presence of CAD, as the study population consisted exclusively of patients undergoing investigations for this condition. HRQoL is often conditioned on disease stage, age and sex across studies. Two studies[109, 115] conditioned HRQoL on recurrence of cancer, with one of the studies [115] further conditioning these outcomes on histology, type of treatment and response, and time post-treatment.

The costs of patients with lung cancer seem to mostly reflect immediate surgical treatment upon diagnosis. One study[114] also considers palliative treatment for some patients. Only one study[118] conditions the costs of treatment on disease stage; this study also applies a cost penalty to patients with delayed diagnosis (false negatives) consisting of the cost of one GP appointment and one additional CT scan.

The outcomes of patients with benign SPNs are less well described in the publications. Where described survival and HRQoL is mostly conditional on age and sex, and reflect the outcomes of the

general population. Costs beyond those accrued in the diagnostic pathway are not reported or included for patients with benign nodules.

The only study which uses UK-specific evidence sources is Rickets et al., 2020[118], the only identified UK study. This study used, as main sources of information on outcomes, Cancer Research UK statistics of mortality with treatment by disease stage at diagnosis and ONS data on other cause mortality, [228] [ONS 2018] estimates of HRQoL from Sturza et al., 2010 [229] and disease costs by stage of cancer also from Cancer Research UK[230].

11.8.1.3 Other value components

Overview of how treatment of benign nodules (true negatives and false positives) has been considered

Two models are not explicit about allowing false positive results [113, 114] but assume that some benign nodules that were identified as true negatives receive surgical treatment, with morbidity, mortality and cost implications. The proportion true negatives undergoing surgical treatment is defined by assumptions (e.g., on the growth rate of benign tumours [113, 114]) or by the strategy (e.g., in Deppen et al., 2014, [113] when all patients are tested with VATS [assumed a perfect test], all benign nodules receive wedge resection). Both studies consider a probability of benign nodules (0.10) growing at a rate similar to malignant tumours during CT surveillance, which was not supported by robust evidence. These nodules were assumed to be referred to exploratory surgery (with a mortality risk associated), after which they receive no further diagnostic follow-up or treatment.

In four studies,[109, 112, 117, 118] false positive results are allowed in (at least one of) the full diagnostic strategies analysed. This was implemented by considering that all tests in the diagnostic strategy have imperfect specificity. Outcomes i.e., the costs and adverse outcomes of unnecessary treatment, were directly linked to the proportion of false positives derived from the patient flow in the model. Two studies[109, 112]considered imperfect specificity for biopsy, with false positives receiving surgery (wedge resection and lobectomy) resulting in mortality risk and loss of HRQoL (due to diagnostic induced pneumothorax and surgical procedures). Another study [117] considered that false positive results would be followed by wedge resection (using exploratory thoracotomy or VATS), with associated mortality and morbidity risks. These authors applied a life-expectancy deduction to all patients who underwent biopsy and surgical treatments of a duration corresponding to that of the hospital stays due to these procedures. The cost of unnecessary surgical treatment was included in three studies.[109, 112, 117] In two of these studies.[112, 117] the costs of surgery included both the costs of the surgical procedure and of procedural complications. It was not clear if the cost of surgery included the costs of surgical complications in one of the studies.[109] The fourth study[118] did not report how false positives were handled in the model.

None of the models assumed morbidity from surgical treatment would have longer-term consequences.

Overview of how regression of benign nodules was considered

Two studies[115, 117] consider that benign nodules may regress with full resorption potentially leading to earlier discharge from surveillance. The rate of benign nodule regression was based on expert opinion in one study[117] and stated to be parameterised within a natural history model in the other study[115]. However, it is unclear how the consequences of nodule regression in terms of costs, survival and HRQoL were quantified in these models.

11.8.2 Review of cost-effectiveness studies of lung cancer screening

11.8.2.1 Overview of the screening models

This section reports details on the information extracted from the subset of screening models in which the screening review is focussed (see Section 5.3, Table 24). Table 24 summarises the studies, which used a variety of modelling approaches to evaluate the cost-effectiveness of lung cancer screening with LDCT. The complexity of the modelling structure appears to relate to the complexity of the screening regime, with simulation models used to evaluate alternative inclusion criteria for screening (driven by individuals' baseline lung cancer risk: age, smoking status and smoking exposure) and alternative repeated screening regimens. Simpler model structures such as decision-trees, mathematical models (equation based) and other approaches (e.g., actuary models) have been more frequently used to evaluate one-off screening regimes for a population assumed uniform in terms of baseline lung cancer risk.

In this section, we report the indirect value components attributed to screening (i.e. those related to detection) in the identified studies; these are summarised in Table 40.

11.8.2.2 Overview of value components in the lung cancer screening cost-effectiveness models

Table 40 lists the components of value related to classification and the studies in which these were quantified.

Table 40 Screening studies summary: Identification of value components related to detection

Study (year)	True positives				Allows false positives	Other value components
	Earlier diagnosis			Overdiagnosis		
	Stage shift	Within stage	Lead time modelled			
Snowsill (2018), Griffin (2020)	Yes	No*	Yes	Yes	Yes	No
Marshall (2000)	Yes	No	No	Yes	Yes	No

Study (year)	True positives				Allows false positives	Other value components
	Earlier diagnosis			Overdiagnosis		
	Stage shift	Within stage	Lead time modelled			
Marshall (2001)						
Yang (2017)	Yes	No	Yes	Yes	Yes	Radiation exposure
Pyenson (2012)	Yes	No	Yes	Yes	NR	No
Pyenson (2014)						
Villanti (2013)				No		
Ten Haaf (2017)	Yes	No	Yes	Yes	Yes	No
Tomonaga (2017)						
Toumazis(2018)	Yes	No	Yes	Yes	Yes	No
Whynes (2008)	No	NA	Yes	No	Yes	No
Field (2016), Field (2016a)	Yes	No	Yes	No	Yes	No
Hinde (2018)						No
Hofer (2018)	Yes	No	Yes	No	Yes	Early recall

* The authors quantified earlier detection within stage but this was not modelled to impact on outcomes; NA, not applicable

All studies model the impact of early diagnosis, [94, 125 137;128, 138, 140, 141, 149;129, 133, 134; 144-146, 150, 152] and all but one [150] model stage shift with early diagnosis as part of the mechanism of value. Within-stage early diagnosis happens when a screen-detected tumour is at the same stage as it would have been if detected clinically. One study[140, 141] discusses that it is possible to accrue a survival benefit from early within-stage diagnosis, but this study did not link early within stage diagnosis to survival outcomes due to constraints on mortality in the model structure. The majority of studies explicitly modelled lead time. [94, 125, 128, 129, 137, 138, 140, 141, 144-146, 149, 150]

Overdiagnosis in the context of the screening models is defined as the increased detection with screening of tumours that would not have been clinically detected, and, therefore are not assumed to have a survival benefit from treatment. Some studies consider overdiagnosis in base-case and/or scenario analyses. [137, 138, 140, 141, 144-146]

Most studies allowed false positive results to screening[94, 125, 140, 141, 144-146, 150;128, 129, 133, 134, 152]

One study[129] considered earlier recalls, by assuming that a proportion of individuals who screened positive would not be referred immediately to a pulmonologist, but rather receive an early recall CT scan three to six months after the screening scan which had identified the nodule as suspicious. This could impose additional delays shortening the time interval between screen detection and diagnosis.

Finally, Yang et al., 2017, [152] considers the impact of radiation exposure.

11.8.2.3 Evidence linkage for earlier/increased detection of lung cancer

Table 41 summarises in further detail the evidence linkage regarding earlier detection of lung cancer. Where there is increased detection with screening strategies compared to no screening, this may result mostly in overdiagnosis, rather than on more patients receiving early treatment that translates into survival gains. Thus, the modelling of overdiagnosis is reported in this section alongside that of earlier diagnosis.

Table 41 Modelling of value components relating to earlier/increased detection of lung cancer

Study (year)	Pre-clinical to clinical progression modelling				Disease stage at diagnosis, if not informed by the progression model	Overdiagnosis
	Structural assumptions on progression model				Mechanism	
	Stages in sequence?	Clinical progression modelled?	Individual heterogeneity modelled?	Other		
Snowsill (2018), Griffin (2020)	Yes	No	Yes	--	Informed by clinical to pre-clinical progression model	Model output that results from simulating the natural history of the disease and screening accuracy.
Marshall (2000)	Not modelled				Informed directly by the effectiveness data	No evidence linkage – modelled directly on outcomes
Marshall (2001)						
Yang (2017)	Not modelled				Informed directly by the effectiveness data	Not modelled
Pyenson (2012)	Not modelled				Informed directly by the effectiveness data	Scenario analysis assuming 5% or 20% more individuals on stage A without any reduction of patients in stage B and C.
Pyenson (2014)						
Vilanti (2013)						Not modelled
Ten Haaf (2017)	Yes	No	Yes	--	Informed by the pre-clinical to clinical progression model	Model output that results from simulating the natural history of the disease and screening accuracy.
Tomonaga (2017)						
Toumazis (2019)	Yes	No	No	--	Informed by the pre-clinical to clinical progression model	
Whynes (2008)	Not modelled				Disease stage at diagnosis not modelled. The impact of early diagnosis is captured directly on survival without modelling the shift	Not modelled
Field 2016, Field 2016a	Not modelled				Informed directly by the effectiveness data	Not modelled
Hinde (2018)						
Hofer (2018)	Yes	Yes	No	--	Informed by clinical to pre-clinical progression model	Not modelled

Models with pre-clinical to clinical progression component

Studies that explicitly modelled disease progression from pre-clinical to clinical presentation did so by using a natural disease history model. Usually, a natural history model is informed by a set of observed transition probabilities estimated from relevant clinical studies. Pre-clinical transition probabilities between disease stages are not observable, and, therefore these probabilities cannot be directly informed by comparative evidence from RCTs. One potential way to estimate these unobservable probabilities is to use calibration methods, which allow comparing plausible model outputs to empirical data (such as disease incidence, stage distribution of cancer by type of deselection, lung cancer mortality rates, etc.), the calibration targets, and vary model inputs to establish the parameter values that best fit the data.[231] These models [129, 140, 141, 144-146] apply calibration methods to infer preclinical progression probabilities. For example, Ten Haaf et al., 2017[145] uses comparative evidence from two screening trials (NLST and PLCO) on observed stage distribution at diagnosis and number of cancers detected by intervention arm and type of detection to calibrate stage distribution at diagnosis (combined with evidence on cancer incidence and survival from other sources) to estimate progression probabilities amongst other parameters (e.g., screening diagnostic accuracy). Preclinical to clinical progression probabilities have also been estimated by calibration based on observational rather than experimental evidence. For example, in one study [129] the preclinical to clinical probabilities were calibrated using observational data on incidence and observed stage distribution in cancer patients not exposed to screening (while diagnostic accuracy was sourced from a separate simulation study[156]). It is not clear how preclinical to clinical progression was informed by these data, given the apparent lack of data on screened patients.

The four models track the movement of individuals over time across preclinical disease stages until they are detected either by screening or clinical presentation (using a patient-level simulation [140, 141, 144-146] or a cohort approach[129]). These models considered stage specific preclinical to clinical progression probabilities, and two models further conditioned these probabilities on tumour histology.[144-146] Although health states differed across models, all imposed a common structural assumption that patients would progress sequentially from less to more advanced disease stages in the pre-clinical model. Only one of the studies modelled progression beyond the point at which disease becomes clinically presenting. [129] Two of the simulation models [140, 141, 146] allow for between individuals heterogeneity.

One model took a different approach to model pre-clinical to clinical progression, which was explicitly based on tumour growth. The natural history model used by Toumazis et al., 2019[146] (described in detail in a separate publication) [232] tracks tumour growth and relates this to pre-clinical to clinical progression (and also probability of treatment being curative). The model assumes an exponential growth function for the primary tumour (parameterised with VDT) and a tumour size

threshold (VC) before which detection and treatment of the primary tumour is assumed to be curative. If the tumour is not treated before reaching this threshold, the lethal metastatic burden starts to increase exponentially as a function of the size of the primary tumour size. At a certain lethal metastatic burden threshold (k1) metastases become observable and patients whose disease is detected after this threshold are assumed to have advanced stage disease. In the model cancer can be clinically detected due to either the primary tumour or metastasis, dependent on which prompts detection present first. Cancer can be clinically detected when the primary tumour reaches second size threshold (VP) or a second lethal metastatic burden threshold (k2). The lethal metastatic burden thresholds are both defined as a fraction of maximal metastatic tolerance level (BD), where BD represents the point at which metastases become the cause of death.

In these models the lead time and stage shift between screened and clinically detected cancers is informed by the tracking preclinical to clinical disease progression combined with the screening accuracy.

In the simulation models overdiagnosis was modelled as an output by quantifying the proportion of tumours that are detected with screening in excess of those clinically presenting with a no screening strategy.[140, 141];[144-146]

Models without pre-clinical to clinical progression component

Models without a natural history model component [94, 125 128, 133, 134, 137, 138, 149, 150, 152] did not model pre-clinical to clinical progression and, with the exception of the study by Whynes et al., 2008, linked effectiveness data on stage distribution combined with assumptions on lead time, to survival outcomes.

Four of studies (Whynes, 2008; Field, 2016; Field 2016a; Hinde, 2018) apply a common methodological approach, which uses lifetables capturing general population and cancer-specific mortality (for patients with i) screen detected and ii) clinically detected cancer) to estimate survival benefits associated with earlier diagnosis. This approach assumes a common general population mortality rate for screened and unscreened (clinically detected) patients up to the assumed age of detection with screening, at which point survival diverges between the two populations. The survival function of screened patients beyond the age of detection with screening follows a negative exponential model that implies an increased mortality rate from the age of detection. The age of clinical detection is estimated by adding an assumed lead time to the age of detection with screening. The survival of patients with clinically detected cancer is assumed to follow general population mortality until detection, with a negative exponential model fitted beyond that point. Both the clinically and screen detected population mortality rates become the same as the general population at the point (beyond detection) where the mortality rate predicted by each of the exponential functions

exceeds that of the general population. In the original mathematical model developed by Whines et al., 2008, survival estimates were not conditioned on disease stage at detection (only age); the survival of screened patients beyond detection was directly informed by the ELCAP study (1- and 10-year survival rates in the screening arm). This model assumed a homogeneous cohort of male patients, and a single lead time estimate for the cohort. The other two studies[94, 125, 128] adapted the original model so as to condition survival on disease stage and age at screen detection (as well as sex). In Field et al., 2016 [94, 125] the survival model is solved for each cancer screen detected in the UKLS trial (authors describe this as a simulation), using life tables specific to the patient's sex and age at screening, and stage specific post-detection mortality. The stage specific post-detection mortality was informed by the ELCAP study (as UKLS did not have sufficient follow-up data) for the screen-detected population, and from UK cancer statistics for the clinically detected population. Patients with stage 4 disease at screening were assumed to have no survival benefit from screening. The stage distribution at diagnosis was sourced from UK screening pilot trials data for screen detected cancers (UKLS in Field et al., 2016,[94, 125] and the Manchester lung cancer screening pilot for Hinde et al., 2018[128]), and by UK cancer statistics for clinically detected cancer.

The approach taken to reflect stage-shift in two other models[137, 138, 149];[133, 134] sourced the stage distribution for the 'no screening' strategy from registry data (SEER), while for the screening strategies this was sourced from the screening arm of RCTs.

Yang et al., 2017, [152] assumed that stage (and histological) distributions of screen-detected and non-screen-detected lung cancers in the screening and 'no screening' strategies were the same as those for CT-screening and radiography-screening in the NLST, respectively.

Overdiagnosis in these models was considered variability; one of the models [137, 138] an additional proportion of individuals on stage A (5% or 20% in each of the scenario analyses) for the screening strategy compared to base-case without any reduction of patients in stage B and C. The authors did not justify the range of values tested in this scenario analysis. Another model [133, 134] explored the impact of potential overdiagnosis directly on outcomes, without establishing a link between an estimate of overdiagnosis and outcomes.

Handling biases arising from early detection of lung cancer

There are two common type of biases that can affect the estimation of survival benefits of patients with screen detected cancer: i) lead time bias, and ii) length time biases. [233]

Lead time bias arises from screening prolonging the interval between diagnosis and death, even if early treatment had no effect on patient survival, as diagnosis occurs earlier with screening compared

with clinical detection. [233, 234] Therefore, when quantifying the survival benefit attributable screening, the lead time as to be excluded from the survival gains of screened vs. unscreened patients.

Another type of bias, length bias, may arise due slow growing tumours being more likely to be detected by screening, given the interval between screening appointments. In contrast, fast growing tumours will progress quickly from preclinical to clinical stages and will be more likely to be clinically detected. Since slower growing tumours usually have better prognosis, the survival benefit of screened patients could be driven by the identification of proportionally more of the less aggressive, slow-growing tumours.

Depending on how effectiveness data is used to parameterise each model, adjustments may be needed to ensure that these biases are not introduced.

In the actuary model[137, 138, 149] lead time was assumed to have an homogenous duration (two or three years) and this estimate was deducted from the survival gains predicted for patients with screen detected cancers. Some studies[133, 134] handled lead time bias in scenario analysis only; the adjustment was limited to the deduction of one year from the survival gains of the screened patients. The lead time duration assumption in these studies[133, 134, 137, 138, 149] was not justified.

Whynes et al., 2008, [150] also assumed a single lead time estimate for screen-detected tumours (8 years), which is stated to correspond to the upper bound of the range values described in screening trials literature. Other studies [150];[94, 125, 128] assumed stage specific lead time estimates. These were informed by assumptions: the double of the difference between mean subject ages at screen detection by stage and the ages of symptomatic presentation currently observed in the UK was assumed for cancers detected by screening at stages 1-3; stage 4 cancers detected by screening at stage 4 disease at screening were assumed to have no lead time. Lead time was used in these models to determine age at screen detection, point at which the survival model for patients with screen detected tumours start following a different survival model.

Yang et al., 2017, [152] used a differences-in-differences methodology to deal with lead time bias in their model. The differences in expected life years lost due to cancer conditional on stage between screened vs. unscreened patients were estimated against a reference age- and sex-matched population to adjust for age at diagnosis. By estimating the survival estimates for screened and unscreened patients relative to the reference population for each group of patients instead of directly against each other, the model does not incorporate the difference in age at diagnosis between groups as a survival benefit for the screened patients group.

Models that simulate the natural lung cancer history with a pre-clinical to clinical progression component, do not require assumptions on the duration of lead time, as lead-time is estimated by the

model as an output. Lead time bias may still be incorporated if structural assumptions on mortality allow for survival benefits of the screened patient to stem (partly) from early diagnosis alone. Only one of the simulation models states how leading time bias was handled. [140, 141] This model assumed the same survival for lung cancer in each stage, regardless of detection type (screen detected or clinically presenting). Furthermore, the age of lung cancer mortality was assumed to not be brought forward by screening. This imposed a lower bound on survival of A + B, where A represents the expected survival in the later stage (in which the cancer would have presented absent screening) and B is the lead time.

Only one of the identified models reported handling of length bias. [140, 141] The authors address this bias via the same survival constraint that is used to handled lead time bias.

Modelling of link between disease status and staging, and outcomes

Table 42 summarises how the link between disease status and staging was established in the identified studies.

Table 42 Modelling of link between disease status and staging, and outcomes - screening studies summary

Study (year)	Link to outcomes	Staging categorisation	Outcomes	Disease status						
				Lung cancer				No Lung cancer		
				Conditional on	Assumptions	Overdiagnosis	UK relevant source	Conditional on	Assumptions	UK relevant source
Snowsill (2018), Griffin (2020)	Direct link with staging	IA, IB, IIA, IIB, IIIA, IIIB, IV	Survival	Staging, age	. Handling of lead time bias - Same survival for lung cancer in each stage regardless of type of detection (screen vs clinical) . Lung cancer mortality at pre-clinical stages – No . Other:	Constraint on survival by type of detection	No	Age, sex, smoking	--	ONS, [235] [236] Institute and Faculty of Actuaries [237]
			HRQoL	Staging, screening	Constant with time	--	No	Smoking, age, sex, FP result, screening	Constant	Health Survey for England, 2014 [238]
			Costs	Staging, time post diagnosis, FN result, EoL	Time varying	--	McGuire 2015; [239] Round 2015; [240] Kennedy 2016 [241]	NA	--	--
Marshall (2000), Marshall (2001)	Direct link with staging	I, II, IIIA, IIIB, IV	Survival	Staging, tumour size (stage I), age, sex	Handling of lead time bias – 1-year adjustment in	Scenario reducing survival benefit for patients with screen	No	Age, sex, race	--	No

Study (year)	Link to outcomes	Staging categorisation	Outcomes	Disease status						
				Lung cancer				No Lung cancer		
				Conditional on	Assumptions	Overdiagnosis	UK relevant source	Conditional on	Assumptions	UK relevant source
					scenario analysis Lung cancer mortality at pre-clinical stages – No Other: Same survival for lung cancer in each stage regardless of type of detection (screen vs clinical)	detected cancer by 1 year				
			HRQoL	Staging	Constant with time	--	No	Sex, smoking	Constant	No
			Costs	Staging	Constant with time	--	No	NA	Constant	No
Yang (2017)	Direct link with staging	I, II, IIIA, IIIB, IV	Survival	Staging, histology	Handling of lead time bias – Differences-in-differences approach Lung cancer mortality at pre-clinical stages - NA		No	Age, sex	Time varying with age	No
			HRQoL	NR	NR		No	Age, sex	Time varying with age	No

Study (year)	Link to outcomes	Staging categorisation	Outcomes	Disease status						
				Lung cancer				No Lung cancer		
				Conditional on	Assumptions	Overdiagnosis	UK relevant source	Conditional on	Assumptions	UK relevant source
			Costs	Staging, histology, radiation exposure	Constant		No	NA	--	--
Pyenson (2012), Pyenson (2014) Vilanti (2013)	Direct link with staging	A, B, C	Survival	Staging, age, sex	Handling of lead time bias – lead time offset used to correct survival estimates Lung cancer mortality at pre-clinical stages - NA	Additional patients assumed to be overdiagnosed have the same survival as Stage A patients.	No	Age, sex	---	No
			HRQoL*	Staging, age, sex	Time varying with age	NA	No	Age, sex	Time varying with age	No
			Costs	Staging, time post diagnosis	Time varying – becomes constant from year 5 onward	Additional patients assumed to be overdiagnosed have the same costs as Stage A patients.	No	NA	--	--
Ten Haaf (2017)	Direct link with staging	IA, IB, II, IIIA, IIIB, IV	Survival	Staging, histology, sex, detection type (chest radiography vs LDCT screening)	Handling of lead time bias – No Mortality at pre-clinical stages - No		No	Birth year, sex, smoking history	Time varying with age	No

Study (year)	Link to outcomes	Staging categorisation	Outcomes	Disease status						
				Lung cancer				No Lung cancer		
				Conditional on	Assumptions	Overdiagnosis	UK relevant source	Conditional on	Assumptions	UK relevant source
Tomonaga (2017)			HRQoL	NA	--	--	--	NA	--	--
			Costs	Staging**, age, sex, phase of care (initial, continuing, terminal care)	Time varying – by phase of care		No	NA	--	--
Toumazis (2019)	Direct link with staging		Survival	Staging, histology, sex, cure (via tumour size and metastatic burden at detection)	Handling of lead time bias – No Lung cancer mortality at pre-clinical stages - No		No	Birth year, sex, smoking history	--	No
			HRQoL	Age, sex, staging, detection type (clinical or screening) and histology, treatment, time post successful treatment, EoL	Time varying - Lung cancer survivors after 5 years post primary diagnosis with no further cancer events return to normal health-states utilities		No	Age, sex	Time varying with age	No
			Costs	Type of cancer treatment, phase of care (initial,	Time varying - according to cancer care phase		No	NA	--	--

Study (year)	Link to outcomes	Staging categorisation	Outcomes	Disease status						
				Lung cancer				No Lung cancer		
				Conditional on	Assumptions	Overdiagnosis	UK relevant source	Conditional on	Assumptions	UK relevant source
				continuing, terminal care)						
Whynes (2008)	Direct link between type of detection and outcomes	NA	Survival	Age, detection type	Handling of lead time bias – No Lung cancer mortality at pre-clinical stages – NA	--	No	Age	--	Lifetables from Government Actuary Department[242]
			HRQoL	Detection type	Constant - single utility adjustment for clinically presenting cases	--	NA	--		
			Costs	Timing of cancer treatment (early vs later)	Constant	--		NA--		
Field (2016, 2016a) Hinde (2018)	Direct link with staging	I, II, III, III, IV	Survival	Staging, age, sex, detection type	Handling of lead time bias – No Lung cancer mortality at pre-clinical stages – NA	--	UK cancer survival statistics [243-245] for clinically detected	Age, sex	--	Not referenced
			HRQoL	Detection type, age at death	Constant	--	?	NA	--	--
			Costs	Staging, timing of	Constant	--	Field:	NA	--	--

Study (year)	Link to outcomes	Staging categorisation	Outcomes	Disease status						
				Lung cancer				No Lung cancer		
				Conditional on	Assumptions	Overdiagnosis	UK relevant source	Conditional on	Assumptions	UK relevant source
				cancer treatment (early vs later)			Estimated within study Hinde: Cancer research UK [230]			
Hofer (2018)	Mediated via treatment	I, II, IIIa, IIIb, IV	Survival	Staging, treatment type, post-detection stage	Handling of lead time bias – None Lung cancer mortality at pre-clinical stages – Yes Other: Treatment type| stage; post-detection stage| treatment type, surviving treatment	--	No	Age(unclear), smoking	--	--
			HRQoL	Treatment type/ post-detection/treatment health state	Constant in time Same utility on all pre-clinical and no disease stage	--	No	Age	--	--
			Costs	Treatment type	Surviving diagnosed patients not	--	No	--	--	No

Study (year)	Link to outcomes	Staging categorisation	Outcomes	Disease status						
				Lung cancer				No Lung cancer		
				Conditional on	Assumptions	Overdiagnosis	UK relevant source	Conditional on	Assumptions	UK relevant source
					undergoing palliative care incur a fixed cost per cycle					

*In Villanti et al., 2013 only, **In Ten Haaf et al., 2017, only; EoL, end of life; FN, false negative

As previously mentioned, the key component of value is stage shift, and, therefore, in the majority of models survival outcomes for patients with lung cancer on stage distribution are conditional on stage distribution; [94, 125, 128, 129, 133, 134, 137, 138, 140, 141, 144-146, 149, 152]some models also condition this on tumour histology. [144-146, 152] Whynes, 2008, [150] does not condition survival on staging; the survival outcomes of screened patients are informed with cumulative survival probabilities from the ELCAP study, while UK cancer statistics inform these outcomes for patients with clinically detected cancer.

The majority of models with a pre-clinical to clinical progression model all assumed that there is no pre-clinical lung cancer mortality.[140, 141, 144-146]

Some models also conditioned survival outcomes of patients with lung cancer on how disease was detected. [94, 125, 128, 144, 145, 150]

In one model[144, 145] this was implemented via the probability of cure which differs by the stage of detection and between computed tomography and chest radiography for stages IA, IB and II. The authors state that this was to account for the large difference in mortality for these stages between the two screening methods, but do not discuss whether this may have led to lead time biases arising. Other studies[94, 125, 128, 150] used different survival models to inform the survival outcomes of patients according to whether lung cancer was clinically or screen detected. Some studies explicitly state that survival by cancer stage was assumed the same regardless of how cancer was detected (Marshalls, snowsill, griffin). For one of the models, this assumption was made to limit the impact of biases. [140, 141] The authors considered that evidence from screening trials suggesting survival is higher for screen-detected cancers than non-screen-detected cancers, (including those of the same stage) may be partially driven by lead time, length and overdiagnosis biases (see sections 0 and 0).

Two models condition the survival outcomes of patients with lung cancer on nodule size. [133, 134, 146] The natural disease history model by Toumazis et al., 2019,[146]tracks tumour growth and conditions the probability of cure on tumour size at detection and metastatic burden (which is also a function of tumour size). The model assumes that cured patients (treated before the tumour reaches a certain size) can die of other causes, but not due to lung cancer. The model by Marshal et al. [133, 134]stratify lung cancer survival by tumour size (≤10 mm, 11–20 mm, 21–45 mm, >45 mm), for patients with stage I (in addition to stage, sex and age), but this seems to be equivalent to using additional substages within the disease classification (e.g., Ia, Ib, etc.). These studies[133, 134] do not appear to explicitly model tumour growth over the time horizon.

Staging was also linked to HRQoL and/or costs in some models. Some models considered stage specific HRQoL estimates for patients with lung cancer; HRQoL estimates could be constant over time [133, 134, 140, 141]or time varying i) with age [149] or ii) assuming general population utility after 5 years

disease free. [146] Stage specific lung cancer costs were considered in seven models.[94, 125, 128, 133, 134, 137, 138, 140, 141, 145, 146, 149, 152]Of these studies, time varying costs were considered in four models; [137, 138, 140, 141, 145, 146, 149] this was dependent on time elapsed post-diagnosis/treatment [137, 138, 140, 141, 149] and/or phase of treatment. [140, 141, 145, 146]. Two UK models[94, 125, 128] condition costs on the type of detection, with costs of investigation and treatment differing between screen and clinically detected lung cancers. The time point at which these costs are assumed to take place also varies by type of detection according to assumed stage specific lead time (see Section 0). However, not all patients who would have presented clinically will incur investigation and treatment cost, as a stage specific proportion of patients is assumed to die before clinical presentation.

One model[140, 141] considered a temporary (two weeks) disutility from screening based on EQ-5D VAS data from the NELSON trial, which aims to capture anxiety associated with undergoing the intervention.

The survival of individuals without lung cancer was conditioned in most models on age/birth year and sex. A few models[140, 141, 144-146] also considered a reduction in survival due to smoking status, exposure or history. The HRQoL of these individuals was also conditioned on age/birth year [140, 141, 146, 149, 152] and sex [133, 134, 140, 141, 146, 152] across studies. The costs of individuals without lung cancer are not included in any of the models (other than the costs of screening and any further investigations).

A limited amount of UK relevant data sources were identified across the studies. One model informed the survival of patients with clinically detected lung cancer with UK cancer survival statistics by disease stage. (Fields et al.,2016[94, 125] by Walters et al., 2013, [243] and Solomon et al, 2013; [244] Hinde et al 2018, [128] by ONS data). Costs avoided by treating screen detected lung cancer compared to clinically detected cancer were sourced from a Cancer Research UK study [230] in the study by Hinde et al., 2018. [128] The same costs were estimated within the Fields et al., 2016 study[94, 125] with assumptions on resource use informed by National Lung Cancer audit data[246] combined with NHS reference costs (unit costs). Another model[140, 141] based hospital costs of treating lung cancer by stage on the resource use estimates of two English studies (one to inform the first year of treatment [241] and the other for costs beyond first year [239] and the costs of end of life care in an England and Wales modelling study. [240] The studies did not use UK specific HRQoL evidence to inform the outcomes of patients with lung cancer.

For individuals without lung cancer, survival data was informed by UK lifetable data. One model[140, 141]was informed by ONS data [235] adjusted for the risk of lung cancer in smokers, [236, 237] so as to reflect other cause mortality. Whynes, 2008, [150] sourced general population mortality from a Government Actuary's Department source.[242] One model [140, 141]applied a UK specific utility detriment [238]to reflect

the HRQoL of smokers, which was estimated based on evidence from the Health Survey for England 2014.

11.8.2.4 Other value components

Overview of how overdiagnosis/overtreatment has been considered

As mentioned above, the patient-level models with a preclinical to clinical progression component output the number of overdiagnosed tumours. [140, 141, 144-146] These tumours appear to be handled similarly to the other true positives; with the same outcomes associated with stage shift, and the costs, morbidity and mortality associated to further diagnostic investigations. Only one model[140, 141] constrained the survival of lung cancer patients, so that stage specific survival does not vary between screen and clinically detected. This could have mitigated the impact of overdiagnosis by reducing the survival benefit attributed to overdiagnosed tumours.

Some models handle overdiagnosis by relying on assumptions. One model[137];[138] assumed in scenario analysis that 5% or 20% more patients were detected in stage A while maintaining the proportion of patients in the remaining disease states constant, and that the costs and survival outcomes of these additional patients would be equivalent to all other stage A patients. This can be considered a change to disease prevalence (as these additional lung cancers will be 'removed' from the population without the disease). Another study[133, 134] used the scenario analysis reducing the survival benefit of screened lung cancer patients by one year (also used to explore the impact of lead time) to have a sense of the impact of overdiagnosis. None of the assumptions on overdiagnosis explored by these authors[133, 134, 137, 138]in the scenario analyses were supported by empirical evidence.

Overview of how false positives results have been considered

The majority of studies which explicitly modelled false positive results to screening, [94, 125, 128, 137, 140, 141];[129, 133, 134, 144-146, 150]seem to have reflected this as a cost impact due to further unnecessary investigations. Only two models[129, 140, 141] explicitly linked false positives to survival to reflect the disutility associated with subsequent diagnostic follow-up and another to the associated mortality.[146] None of the studies states that false positives receive cancer treatment, although only a few studies [133, 134, 140, 141] explicitly assert that false positives do not receive treatment.

Overview of how early recalls have been considered

As mentioned above, Hofer et al., 2018, [129] considered early recall CT scans for a proportion of patients who screened positive. However, the additional delay between screening and diagnosis for patients with lung cancer does not seem to have been modelled, and impact seem to be reflected only on the cost of the additional scan included for individuals placed on early recall.

Overview of how radiation exposure has been considered

In Yang et al., 2017,[152] the impact of radiation exposure was applied as lifetime cost to capture the health care costs of patients who die from radiation induced cancer. However, it was unclear to whom did this impact apply and how did radiation exposure differ across strategies.