Evidence overview: EarlyCDT Lung for lung cancer risk classification of solid pulmonary nodules

This overview summarises the main issues the diagnostics advisory committee needs to consider. It should be read together with the final scope and the diagnostics assessment report.

1 Aims and scope

EarlyCDT Lung is a blood test that could be used alongside standard care to assess the malignancy risk of solid pulmonary nodules found by chest CT or X-ray. Incidental findings of pulmonary nodules in asymptomatic individuals are an increasingly common clinical dilemma encountered by lung cancer clinicians. EarlyCDT Lung could help identify malignant pulmonary nodules that need immediate treatment or a biopsy. This could result in treatment being offered earlier, potentially giving improved patient outcomes. It could also reduce the number of people on CT surveillance, patient waiting times and radiologist time, enabling efficient use of NHS resources.

The EarlyCDT Lung test is an enzyme-linked immunosorbent assay method. It is manufactured by Oncimmune and is available as a CE-IVD kit. EarlyCDT Lung 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). Each test kit contains reagents used for the measurement of autoantibodies in up to 10 patient samples (depending on usage). The result can be interpreted by laboratory staff and positive results would be reviewed by the lung cancer multidisciplinary team.

A blood sample is positive when at least 1 of the 7 autoantibodies detected by the test is elevated above a predetermined cut-off threshold. There are 2 cut-off thresholds recommended by the company for each autoantibody, and the test can give 2 positive results: positive moderate or positive high (with

different accuracy to detect malignancy). The thresholds were set to give a high specificity with the aim of detecting nodules that are likely to be malignant (a rule-in test). A positive EarlyCDT Lung test result updates a person's pre-test risk to help clinicians make decisions about further testing or intervention. The pre-test risk is unchanged if the EarlyCDT result is negative.

For nodules over 8 mm, in current NHS practice the Brock model is used to calculate pre-test risk of malignancy, which is based on the patient's gender, age, smoking history, and other risk factors. The Herder model is used to calculate malignancy risk following a Brock risk assessment over 10% and a positron emission tomography CT (PET-CT) scan. The Herder model is based on patient characteristics, nodule characteristics, and the degree of F-fluorodeoxyglucose uptake on PET-CT.

The pre-test malignancy risk scores are updated by positive-high and positive-moderate EarlyCDT Lung results. The calculation of post-test malignancy risk from the pre-test risk and the EarlyCDT Lung test result is described in Healey et al. (2017), and uses accuracy estimates for the EarlyCDT Lung reported in that study.

The aim of the assessment was to review existing evidence on the potential clinical and cost effectiveness of the EarlyCDT Lung test for lung cancer risk classification of solid pulmonary nodules. The external assessment group (EAG) based the use of EarlyCDT Lung in the context of the British Thoracic Society (BTS) pathway (see figure 2 in the final scope). The guidelines recommend the use of different interventions based on an estimated risk of malignancy (which the EarlyCDT Lung can adjust). In summary, for people with risk scores below 10% risk, CT surveillance is generally used, and above 70% risk referral for excision or non-surgical treatment is generally used. Within the intermediate group (between 10% and 70% risk of malignancy) the guidelines for subsequent care are more varied, with possible management options including image-guided biopsy, CT surveillance and excisional biopsy.

Clinical decisions are based on risk of malignancy and additional factors such as patient fitness and preferences, and nodule characteristics.

Following consideration of evidence identified during the scoping stage of the assessment on the use of EarlyCDT Lung in the clinical population defined in the scope, NICE concluded that it would not be beneficial to develop an economic model to assess the cost effectiveness of the technology at this time. We are aware of the clinical interest in this area. We believe that identifying areas for further research and encouraging data collection would be useful to the system and could facilitate an assessment of the cost effectiveness in the future. The final guidance will focus on identifying areas for further research or data collection but will not make recommendations on cost effectiveness at this stage.

Decision question

What is the clinical and cost effectiveness of the EarlyCDT Lung test for the investigation and diagnosis of lung cancer in people with pulmonary nodules?

Populations

People who do not have a previous history of cancer and have pulmonary nodules:

- between 5 mm and 8 mm in diameter or 80 mm^3 to 300 mm^3 in volume
- greater than 8 mm diameter or 300 mm^3 volume plus:
  - a malignancy risk less than 10% using the Brock model, or
  - a malignancy risk between 10% and 70% using the Brock model
  - a malignancy risk less than 10% using the Herder model, or
  - a malignancy risk between 10% and 70% using the Herder model.

If data is available, the following subgroups will be analysed:

- Incidentally detected pulmonary nodules.
- Pulmonary nodules identified through screening.

- Pulmonary nodules in people who have symptoms of lung cancer.

Interventions

EarlyCDT Lung test

Comparators

Current clinical practice in the NHS to assess risk of malignancy without use of EarlyCDT Lung:

- Brock model
- Herder model

Healthcare setting

Secondary care

Further details, including descriptions of the interventions, comparator, care pathway and outcomes, are in the final scope for EarlyCDT Lung.

2 Clinical effectiveness evidence

The external assessment group (EAG) did a systematic review to identify evidence on the clinical effectiveness and diagnostic accuracy of EarlyCDT Lung technology for lung cancer risk classification of solid pulmonary nodules.

Find the full systematic review results on pages 48 to 52 of the diagnostics assessment report.

Overview of included studies

There were 47 references that met the inclusion criteria for the systematic review. These covered 6 distinct populations and included 4 systematic reviews and 1 case report. The systematic reviews included only studies that were also already identified in the EAG's searches and so were not considered further.

None of the populations explicitly followed the relevant British Thoracic Society (BTS) diagnostic pathway, in which nodules are identified by CT scan prior to an EarlyCDT Lung test. The EAG focused on the 5 populations that reported data for patients with pulmonary nodules identified by CT scans. The total sample size of these 5 populations was 695 patients with pulmonary nodules, including 97 diagnosed cancer cases.

The first population included North American patients with Health Insurance Portability and Accountability Act (HIPAA) authorization, who were selected based on clinical judgement as having a high risk of developing lung cancer. The EAG commented that most patients receiving an EarlyCDT Lung test did not have pulmonary nodules, and for those that did, it was unclear whether the nodules were identified before or after EarlyCDT Lung was done. The ability of EarlyCDT Lung to distinguish between malignant and benign nodules was assessed in this population, and results were reported in several publications. Data from the HIPAA population was used to inform the illustrative graph proposed by the company for general use (Healey et al. 2017). The EAG used data reported in Massion et al. (2017) for this population, which it considered to be the most recent and comprehensive paper. This study included 166 people with lung nodules, 35 of whom had diagnosed cancer. Patients were followed up at 6 months, rather than at 1 or 2 years as recommended in the BTS guidelines.

The second population also comprised US patients at high risk of developing lung cancer, as part of an EarlyCDT Lung Cancer Screening (LCS) study (NCT01700257). At the time of this assessment only interim results reported in conference abstracts were available. One study of this population, which investigated CT scan alone compared with EarlyCDT Lung test in addition to CT scan, was included in the EAG's main analyses. The Jett et al. (2017) study included 352 people with nodules and 7 with diagnosed cancers. Currently, the full published results are not available.

As part of the German Lung Cancer Screening Intervention Trial (LUSI), a retrospective, nested case-control study used EarlyCDT Lung to screen stored blood samples collected prior to identification of nodules and cancer diagnosis (Gonzalez et al. 2021). This study included 136 people with suspicious nodules, 46 of whom had diagnosed cancer.

The 2 final populations were from very small studies. First, a US study of 25 patients with indeterminate risk nodules (Lin et al. 2016) and, second, a follow-up study in Hong Kong of 10 patients with lung nodules (Lau et al. 2017). Both were reported only as conference abstracts.

Details of the included studies can be found in table 5 (page 53) of the diagnostics assessment report.

Study quality

The test accuracy data in the HIPAA (Massion et al. 2017) and the German (Gonzalez et al. 2021) studies were assessed for the risk of bias and applicability using the QUADAS 2 tool. The EAG stated that both studies were at high risk of bias due to patient selection, with HIPAA additionally so in the flow and timing domains. The test positioning within the diagnostic pathway (usage, and interpretation of the test), use of a sub-optimal reference standard and divergence in follow-up periods (6 months [Massion et al. 2017] versus 12 or 24 months [BTS guidelines]) lead to concerns about the applicability of these studies to NHS practice.

Since only abstracts were available for the other populations, quality assessment of the other studies was not possible. An overview of the QUADAS 2 assessment is shown in table 6 of the diagnostics assessment report (page 54).

Intermediate outcomes

Diagnostic test accuracy

The reported diagnostic accuracy across the 5 populations was broadly consistent, with specificity of over 90% but low sensitivity of under 30%.

The use of different thresholds in the studies may have impacted on the reported test accuracies. The Mission et al. (2017; HIPAA population) study results used the threshold for the commercial form of EarlyCDT Lung at that date. This resulted in the highest sensitivity (38%; 95% confidence interval [CI] 22 to 54) across the 5 populations, but also the lowest specificity (86%; 95% CI 79 to 92.1).

The German study (Gonzalez et al. 2021) used 2 thresholds: 'high specificity' and 'moderate specificity'. The sensitivity estimates were much lower than for the HIPAA studies (13%; 95% CI 4.9 to 26.3) with no change in sensitivity between thresholds. Specificity dropped from 95.6% (95% CI 89.0 to 98.8) to 91.1% (95% CI 83.2 to 96.1) when using both the positive-high and positive-moderate cut-offs as a positive EarlyCDT Lung result, compared with using just the positive-high cut-off.

The diagnostic accuracy of EarlyCDT Lung testing in patients with nodules were not reported using only the higher specificity cut-off in any HIPAA study. As the test cut-off used was unclear for the other 3 populations that were reported only as conference abstracts, a meta-analysis at specific test cut-offs was not possible.

Table 1 summarises the meta-analyses results of diagnostic accuracy outcomes, using data from the 5 populations: HIPAA (Massion et al. 2017), Hong Kong (Lau et al. 2017), EarlyCDT LCS (Jett et al. 2017), US (Lin et al. 2016) and Germany (Gonzales et al. 2021). Data was not adjusted for possible variation in prevalence across the studies.

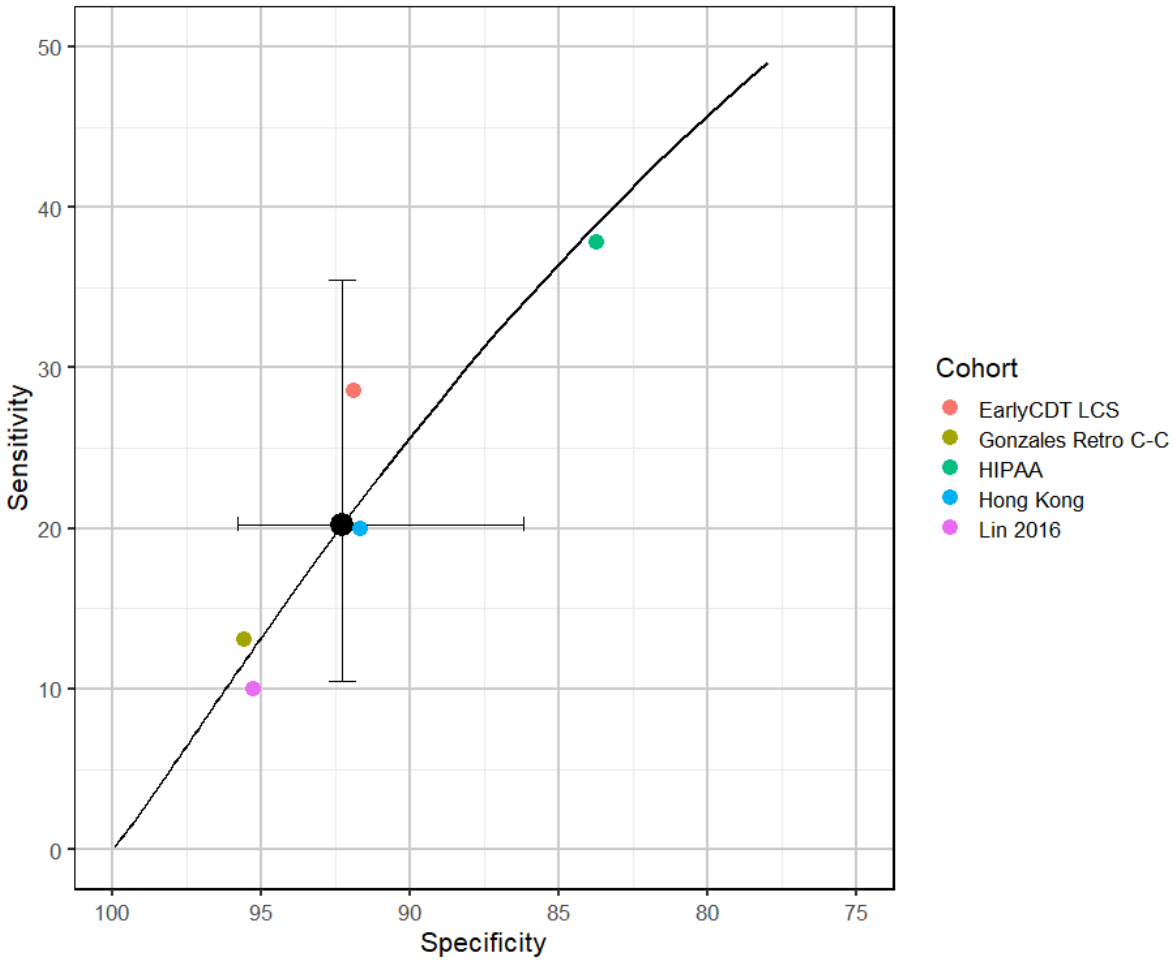
The EAG note that because the studies used different EarlyCDT Lung cut-offs, the summary sensitivity in the bivariate model may not be reliable. Instead, they predict that EarlyCDT Lung has around 26% sensitivity at 90% specificity, or 12% sensitivity at 95% specificity using the HSROC curve (table 1 and figure 1). The area under the HSROC curve was 0.694, suggesting poor to moderate overall diagnostic accuracy.

Table 1. Diagnostic accuracy meta-analyses

	Univariate analysis (95% CI)	**Bivariate analysis (95% CI)**	**HSROC**
Sensitivity	22% (11 to 37)	20.2% (10.5 to 35.5)	26% (at 90% specificity) 12% (at 95% specificity)
Specificity	92% (86 to 95)	92.2% (86.2 to 95.8)	-
PPV	32% (11 to 64)	-	-
NPV	85% (63 to 95)	-	-
DOR	3.32 (1.75 to 6.31)	-	-
AUC	-	-	0.694

Abbreviations: AUC, area under curve; DOR, diagnostics odds ratio; NPV, negative predictive value; PPV, positive predictive value.

Figure 1. EarlyCDT Lung - bivariate meta-analysis and HSROC curve (taken from figure 8 of the diagnostic assessment report)



A more detailed summary is presented in figures 6 to 8, pages 57 to 58 of the diagnostics assessment report.

Healey et al. (2017) accuracy estimates

The EAG commented that the Healey et al. (2017) paper includes a reanalysis of data from a case-control 'optimisation' group within the HIPAA population (reported in Chapman et al. 2012), where cases included people with diagnosed cancer and controls were healthy volunteers (EarlyCDT Lung was performed after cancer diagnosis). The EAG focused on data from Massion et al. (2017), which included people with nodules from the HIPAA population.

The study concluded that there was no significant difference in the diagnostic accuracy between the case-control group and the overall HIPAA population (which included people with and without nodules). Results for the case-control population and estimates derived from other HIPAA populations are presented in table 2. The EAG suggest that the lower sensitivity, specificity and likelihood ratio observed in patients in the HIPAA population with nodules could be driven by poorer diagnostic accuracy among smaller nodules. Smaller nodules were more common in the HIPAA population and are less likely to be present in the case-control group (where cases are confirmed cancer diagnoses). The EAG therefore considered that the diagnostic accuracy estimates from the case-control group may overestimate accuracy in patients with nodules.

Table 2 Diagnostic accuracy as reported in Healey et al. (2017; adapted from table 9 of the diagnostics assessment report)

Group	Sensitivity (95% CI)	Specificity (95% CI)	Positive likelihood ratio (95% CI)
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 mm to 20 mm)	40.0% (15.2 to 64.8)	83.9% (76.2 to 91.6)	2.5 (1.1 to 5.4)
Larger nodules (more than 20 mm)	36.4% (16.3 to 56.5)	91.7% (80.6 to 100)	4.4 (1.0 to 18.4)

Abbreviations: CI, confidence interval; HIPAA, Health Insurance Portability and Accountability Act.

In Healey et al., the diagnostic accuracy for nodule and case-control 'optimisation' groups were claimed to be similar because Fisher exact tests found no evidence of difference (for example, Fisher exact test for specificity: p=0.28). However, the EAG highlighted that the number of people with

nodules was small (111 people) and so lack of evidence could not be equated to no difference. The EAG considered it inappropriate to assume that diagnostic accuracy for the case-control group is generalisable to patients with nodules.

The 'high-specificity' threshold (98% specificity, with 29% sensitivity) and 'low specificity' threshold (49% specificity for 80% sensitivity) established in Healey et al. (2017) and used by the company to calculate post-test risk following a positive EarlyCDT Lung test, were derived from the case-control group. If these thresholds are an overestimation of diagnostic accuracy they will translate to an overestimation of the post-test risk.

The EAG compared the impact of a positive EarlyCDT Lung test result on risk of malignancy (post-test risk) using accuracy estimates from Healey et al. (2017; as used by the company) and accuracy estimates derived from the EAG's analysis (see figure 1). Table 3 describes the sensitivity estimates at 80% and 98% specificity thresholds used in the simulation study.

Table 3. Simulation study inputs

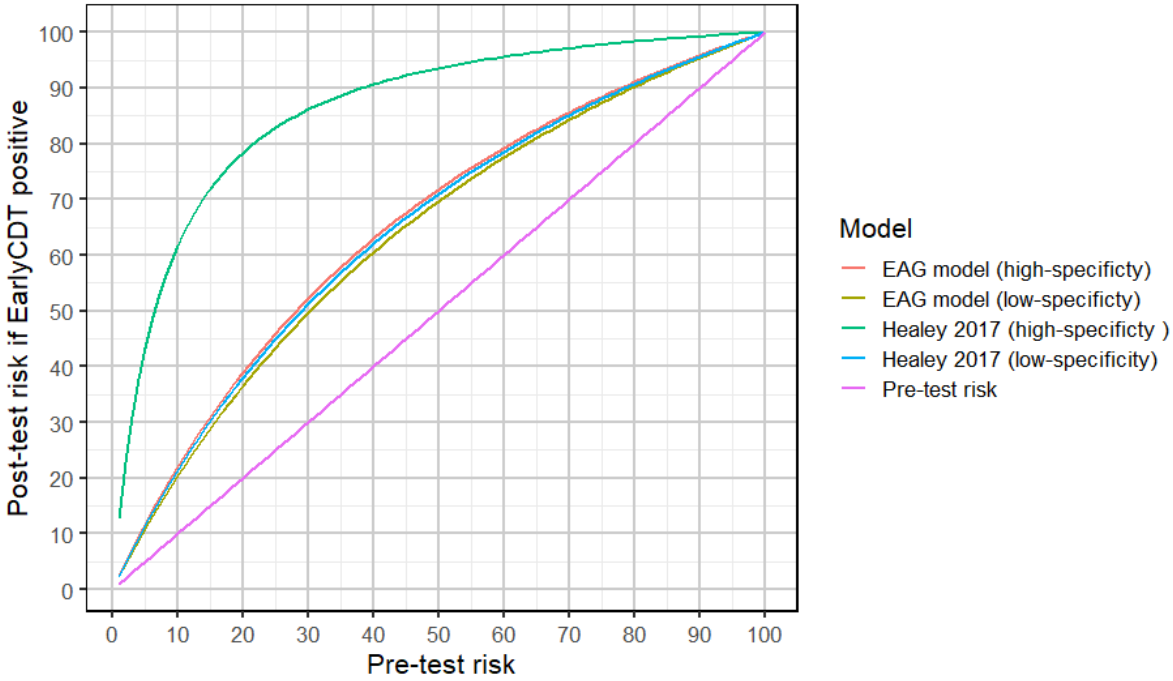
	Sensitivity at high (98%) specificity threshold	**Sensitivity at low (80%) specificity threshold**
Healey et al. (2017)	29	49
EAG meta-analysis	5.1	46

Abbreviations: EAG, external assessment group.

Figure 2 shows the post-test risk following a positive EarlyCDT Lung result estimated using different assumed accuracy of the test (as shown in table 3). This shows that with the EAG estimates there is a smaller increase in risk if EarlyCDT Lung is positive for the 'high-specificity' threshold, because of the much lower predicted sensitivity. If using the EAG's preferred accuracy estimates, rather than those from Healey et al., a positive EarlyCDT Lung test is less likely to change a patient's risk classification to the extent that it would change their care. This is explored by the EAG in a simulation described later.

Further details on the comparison between the Healey et al. model and EAG model is presented on page 66 of the diagnostic assessment report.

Figure 2 Post-test risk using the Healey et al. (2017) and EAG models (taken from figure 10, page 66 of the diagnostics assessment report).



Diagnostic accuracy by nodule size

Diagnostic accuracy was stratified by nodule size in Massion et al. (2017; shown in table 4). The EAG state it is possible that sensitivity declines with increasing nodule size, but specificity increases.

Table 4 Diagnostic accuracy by nodule size

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

Abbreviation: CI, confidence interval.

Combining EarlyCDT Lung with other risk scores

No evidence was identified for combining EarlyCDT with either Brock or Herder models. Massion et al. (2017; HIPAA population) evaluated the Mayo risk model (a functionally similar model to the Herder model) when used on its own and when a positive EarlyCDT test is used to upgrade the Mayo pre-test risk, in comparison with CT imaging. When positive EarlyCDT Lung test results were added to a pre-test risk (Mayo) of 30%, in comparison with using the Mayo risk alone, specificity increased, but sensitivity almost halved. For a 97% specificity threshold, the addition of a positive EarlyCDT Lung result to the Mayo risk increased sensitivity significantly. Further details are on page 59 and summarised in figure 9 of the diagnostics assessment report.

It is unclear how applicable these results are to combination with the Brock or Herder models that are used in clinical practice in the NHS. It should be noted that the 'both positive' approach taken is not what is currently proposed for EarlyCDT Lung, where a positive EarlyCDT result prompts recalculation of malignancy risk.

Comparators

The EAG also reviewed literature on the accuracy of relevant comparators including the Brock and Herder models.

The EAG identified 2 studies that reported complete data on the studied populations: Al-Ameri et al. (2015) for both Brock and Herder, and Perandini et al. (2017) for Herder. The study by Al-Ameri. (2015) included 244 UK people who had a CT scan and 99 confirmed cancer diagnoses. People were referred for PET-CT scan if the nodule was over 5 mm and based on clinical judgement for referral. The Perandini et al. (2017) study in Italy included 259 people and 153 confirmed diagnoses. Both the Brock and Herder risk models were found to have good diagnostic accuracy (area under the curve [AUC] 92%, 95% CI 90 to 95, and AUC 84%, 95% CI 77 to 92, respectively). Although existing data was insufficient to assess accuracy at key risk cut-offs.

More detail on the review and meta-analyses of comparator data can be found in section 3.2, from page 67 of the diagnostic assessment review.

Impact on clinical decision making

As there was a lack of evidence on the diagnostic accuracy and clinical impact of using EarlyCDT Lung in combination with the Brock and Herder risk models, the EAG did a simulation study.

The predicted risks of individuals as assessed by the Brock and Herder models were extracted from figures in Al-Ameri et al. (2015) for both Brock and Herder, and Perandini et al. (2017) for Herder. Brock and Herder risks were analysed separately, as there was no data on the relationship between Brock and Herder risk estimates.

Accuracy data for the EarlyCDT Lung from Healey et al. (2017) and from the EAG's own meta-analysis, were then used to estimate the post-test risks (see figure 2). The EAG's simulation assumes the EarlyCDT Lung test is performed after the Brock or Herder model, as outlined by Healey et al. (2017), for the individuals in the Al-Ameri et al. (2015) and Perandini et al. (2017) studies. The EAG noted this makes a strong assumption that EarlyCDT Lung test results are independent of Brock and Herder risk (given malignancy status). Full detail on the simulation methodology can be found in the diagnostics assessment report from page 82.

Based on the simulation data, the EAG estimated the diagnostic accuracy of Brock alone, Herder alone, Brock with EarlyCDT Lung and Herder with EarlyCDT Lung. The EAG commented that the results show similar ROC curves, suggesting no clear benefit of adding EarlyCDT Lung to either Brock or Herder models. The diagnostic accuracy results for each of these are summarised on pages 84 to 87 in the diagnostic assessment report.

The simulation was also used to estimate the impact of adding the EarlyCDT Lung test to Brock or Herder risk assessment on clinical decision making, in

the populations in the 2 included studies. This was done by estimating the percentages of patients who would be correctly or incorrectly reclassified into higher risk categories following use of the EarlyCDT Lung test. For Brock, this was being reclassified as having above 10% risk. For Herder, 4 initial categories of risk were considered: 0% to 10% (consequential CT surveillance), and 10% to 20%, 20% to 50% and 50% to 70% (the ranges between 10% and 70% were described by the EAG as arbitrary smaller ranges within the intermediate risk range).

Within the low (0% to 10%) pre-test risk groups (Brock model) there were more incorrect reclassifications than correct reclassifications after combining Brock with EarlyCDT Lung (approximately 2:1). This could lead to unnecessary biopsies. The EAG note that it is unclear whether the potential benefits of correctly identifying some malignant nodules will outweigh the harms of these unnecessary biopsies, or what clinical benefits EarlyCDT Lung has on people with smaller nodules where biopsy is not feasible. The increase in risk with a positive EarlyCDT Lung was never sufficient to increase risk to above 70%.

In the intermediate risk group (Herder) above a pre-test risk of 20%, 20% to 35% of patients were correctly reclassified. Many were upgraded into the over 70% post-test risk group. Around 7% of people with benign nodules were wrongly reclassified into a higher risk category. Key results of reclassification in the simulation study are shown as percentages of the pre-test risk group who were reclassified to a higher risk category. The results are summarised in table 17, page 89 of the diagnostics assessment report.

The EAG commented that the variation in results between the Al-Ameri et al. (2015) and Perandini et al. (2017) datasets suggest uncertainty around the exact proportions of patients who will have risk reclassified after EarlyCDT Lung. Due to the lack of evidence on the downstream decisions following risk reclassification, for example the benefits of proceeding directly to surgery,

rather than first receiving a biopsy, the clinical impact of EarlyCDT Lung is uncertain.

EarlyCDT Lung outside the BTS pathway

The Early Detection of Cancer of the Lung Scotland (ECLS) study was considered separately to the 5 cohorts as it did not report data on people with nodules. Instead, the trial assessed the use of EarlyCDT Lung for screening to initiate CT surveillance for earlier diagnosis compared with standard clinical care (symptomatic presentation). The trial screened 12,208 people at risk of developing lung cancer in Scotland.

In the ECLS screening trial, EarlyCDT Lung testing resulted in earlier diagnosis of malignant tumours compared with no screening. The main results were reported in Sullivan et al. (2021). At 2 years, 127 lung cancers were detected, with a higher proportion of stage 3 and 4 cancers detected in the control arm (73%) compared with the intervention arm (59%). The study used positive EarlyCDT Lung results to cause a referral for CT scans, whilst people with a negative result were referred to standard clinical care. This makes it difficult to ascertain if these conclusions would be valid when assessing pulmonary nodules identified within the recommended BTS pathway. More details on pathways outside the BTS guidelines are presented in section 3.1.7 from page 61 of the diagnostic assessment report.

A Danish study was published after the searches for this assessment had been completed. The study by Borg et al. (2021) included 246 patients and 75 were diagnosed with lung cancer. The results were consistent with the EAG's meta-analyses of diagnostic accuracy. The study concluded that EarlyCDT Lung has insufficient sensitivity to be recommended as part of a low-dose CT lung cancer screening programme. The study was excluded as it included people only suspected of having lung cancer and did not report data on people with confirmed nodules.

The EAG did not include a series of case-control studies assessing the potential diagnostic accuracy of EarlyCDT Lung (Boyle et al. 2011 and Lam et al. 2011) in their review, as EarlyCDT Lung was performed after cancer diagnosis, rather than after identification of pulmonary nodules. These studies also used the earlier 6-panel version of EarlyCDT Lung. The EAG provide further explanations of the biases of case-control studies for diagnostic accuracy on page 64 of the diagnostic assessment report.

Clinical outcomes

The EAG reported that, beyond diagnostic accuracy, no data on any of the broader clinical effectiveness outcomes listed in the scope were identified. Therefore, the EAG concludes that there is currently no direct evidence on the clinical value of EarlyCDT Lung when used to assess pulmonary nodules.

Health-related quality of life outcomes

One screening study (ECLS) reported psychological outcomes in 338 people who were EarlyCDT Lung positive. At 3 months and 6 months, there were no statistical differences in lung cancer worry, health anxiety, illness perceptions, lung cancer risk perception or intrusive thoughts, between people with and without pulmonary nodules. No health-related quality of life (HRQoL) outcomes were reported in the target population.

3 Cost-effectiveness evidence

The external assessment group (EAG) did a systematic review to identify any published economic evaluations of EarlyCDT Lung for lung cancer risk classification.

Systematic review of cost-effectiveness evidence

Overview of included studies

The EAG identified 2 studies that both used a decision analytical model approach. Edelsberg et al. (2018) was done in the US, whilst Sutton et al.

(2020) was done in the UK. The positioning of EarlyCDT Lung and target patient population was defined differently in the 2 studies. Edelsberg et al. (2018) assessed people with incidentally detected nodules of 8 mm to 30 mm and intermediate risk (5% to 60%) of lung cancer. Sutton et al. (2020) assessed people with nodules with a size 4 mm to 20 mm, and a risk of lung cancer of 10% to 65%.

Both studies compared EarlyCDT Lung in addition to CT surveillance with CT surveillance alone. A positive EarlyCDT Lung test was assumed to be followed by biopsy (or wedge resection for some people in Edelsberg et al. 2018). People considered to have a benign nodule had CT surveillance until they reach the end of the surveillance period or were considered to have a malignant nodule. In Sutton et al. (2020) nodules were assumed to be biopsied by excision if a person tests positive during CT surveillance. It was unclear how these people were assumed to be managed in Edelsberg et al. (2018).

Both studies used diagnostic accuracy estimates for EarlyCDT Lung from Healey et al. (2017), using either 41% sensitivity and 93% specificity or 98% sensitivity and 28% specificity. The 2 studies are summarised in table 18, page 95 of the diagnostics assessment report.

The incremental cost-effectiveness ratios (ICERs) for EarlyCDT Lung plus CT surveillance compared with CT surveillance alone were about $24,000 per additional quality-adjusted life year (QALY) in the Edelsberg model, and about £2,000 per additional QALY in the Sutton model. Further details of the cost-effectiveness results are on page 99 of the diagnostics assessment report.

The EAG did not consider either of the models suitable to inform the decision problem stated as defined in the NICE scope. It states that they are unlikely to reflect UK clinical practice because:

- Both models used estimates of malignancy prevalence from a study in the US evaluating the use of EarlyCDT Lung that may not be relevant to populations defined in the NICE scope (Tanner et al. 2017). Prevalence estimates are also unlikely to reflect malignancy prevalence at any of the proposed positions for the technology.
- The modelled positioning of EarlyCDT Lung in the diagnostic pathway does not match any of the potential uses of the technology defined in the NICE scope.
- The use of binary test results (positive or negative) in these studies is not appropriate for assessing the additive value of EarlyCDT Lung as part of malignancy risk assessment, as recommended by the company.
- The studied comparators do not include all relevant alternatives.

Further details on the EAG's critique of the relevance of the studies to the NICE scope are described in section 4.4.1 of the diagnostics assessment report, on page 100.

The EAG concluded that existing cost-effectiveness studies cannot directly inform the decision problem. It also commented that neither of the analyses included sensitivity and value of information analyses to an appropriate extent. This limited the relevance of the conclusions reached.

Additional targeted review

The EAG also did 2 further literature reviews of cost-effectiveness modelling studies to support model conceptualisation by identifying different uses of the test that have been modelled as having clinical and economic impact (described as value components by the EAG) and evidence linkage mechanisms. One focused on diagnostic tests or strategies within the diagnostic pathway for pulmonary nodules, and the other on screening strategies for lung cancer. Details on the searches done and results can be found in the diagnostics assessment report, section 5 from page 110.

The EAG commented that existing decision models model the impact of using EarlyCDT Lung on earlier diagnosis of cancer, and consequent improvement in stage at diagnosis (compared with current practice), that relies mostly on assumptions and does not seem to be supported by robust evidence.

The approaches taken in the identified studies were used by the EAG to inform the conceptualisation and analysis of requirements for future modelling.

Economic analysis

Conceptualisation of the decision model and future evidence requirements

Due to the level of uncertainty in evidence, rather than specify a single model structure or approach, the EAG outlined the key evidence requirements and main considerations required to support the development of a future decision analytic model to assess this decision question.

The EAG used influence diagrams to represent the structural relationships between key parameters required for assessing cost effectiveness of EarlyCDT Lung to help conceptualise the model. The analysis identified 3 core components for a future cost-effectiveness assessment of EarlyCDT Lung, which are described in the following sections.

Population

The EAG identified a single study (Al-Ameri et al. 2015) that followed people in the British Thoracic Society (BTS) pathway, and stated that it represented the best current evidence on the UK population on which to base an economic model. But this was a small study (n=186). The EAG suggest that future evidence collection efforts should focus on describing key characteristics that drive value in the population of interest, such as malignancy prevalence, diagnostic or surveillance procedures used, histology and stage distribution at diagnosis.

The EAG highlighted the value of ensuring that future evidence helps understand and describe potential sources of heterogeneity, particularly if this relates to features such as malignancy risk, speed of nodule growth, speed of preclinical progression or long-term health outcomes that are likely to impact cost-effectiveness estimates.

Section 6.2 in the diagnostics assessment report proves further detail.

Defining clinical management decisions

The EAG highlighted that there is uncertainty about which intervention (CT surveillance, biopsy or excision) is offered to people based on their malignancy risk, particularly in the intermediate category (between 10% and 70%). It is therefore difficult to assess the impact of the EarlyCDT Lung in changing risk of malignancy estimates and of how this changes care (and impacts on longer-term outcomes). It also noted that, based on its preferred accuracy estimates identified in the systematic review, the EarlyCDT Lung test may have a limited impact on risk of malignancy estimates. For example, a post-test risk of 70% can only be achieved in individuals with a pre-test risk above 48%.

Based on their analyses, the EAG stated that the EarlyCDT Lung test is unlikely to change referrals to CT surveillance in a number of subgroups, including for individuals with small and low-risk nodules that cannot be biopsied. Groups the test may change management in are people with:

- Low- or intermediate-risk nodules that would have been referred to CT surveillance but that can be biopsied, who could change from CT surveillance to biopsy based on EarlyCDT Lung test.
- Intermediate-risk nodules with a higher pre-test risk score (above 48% according to the EAG's analysis) and that cannot be biopsied, who could change from CT surveillance to excision on the basis of EarlyCDT Lung test.

- Intermediate-risk nodules with a higher pre-test risk score (above 48% according to the EAG's analysis), who could change from biopsy to excision on the basis of EarlyCDT Lung test.

The EAG commented that future evidence should explore the relationship between the risk of malignancy and the likelihood of referral to surveillance, biopsy or excision. Other considerations including patient preference and fitness to receive more invasive tests.

Direct evidence on how EarlyCDT Lung test results affect subsequent management decisions would also support assumptions made about the impact of the test on care. The EAG also noted that future assessment should consider subgroups with restricted care options, for example those who cannot be biopsied, or who are at higher risk of adverse events from this procedure.

Section 6.3 of the diagnostics assessment report provides more detail, including an explanatory influence diagram on page 131, figure 19.

Impact of EarlyCDT Lung arising from changes in management decisions

The EAG considered the short-and long-term impacts on changes to patient care following EarlyCDT Lung testing separately.

Short term impact

These include the impact of escalating management to invasive diagnostics or treatments, including the costs and harms from unnecessary invasive diagnostics or treatments on benign nodules (false positives), and the implications of radiation exposure from increased referral to PET-CT scan.

The EAG highlighted limited evidence on relevant aspects of CT surveillance, biopsy and primary tumour treatment. Sections 6.4.1.1, 6.4.1.2 and 6.4.1.3 in the diagnostics assessment report provides further details.

Overdiagnosis of indolent nodules could occur from use of the test. The EAG commented that 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.

The EAG also commented that further evidence should be generated to provide a better understanding of the rate of resection of benign nodules.

Long-term impact of increased or earlier detection of lung cancer

In existing cost-effectiveness studies, the impact of the EarlyCDT Lung test is assumed to occur through earlier or increased diagnosis of lung cancer. In particular, diagnosis at an earlier stage. This encompasses any health benefits and associated costs of earlier or increased detection of malignant nodules in people with lung cancer (true positive) for whom management is changed. However, the EAG was not aware of any evidence of earlier detection from use of EarlyCDT Lung for incidentally detected nodules. This also applied to increased detection of lung cancer.

In the absence of data on the direct impact of EarlyCDT Lung on the stage at which lung cancer is diagnosed (stage shift), and on the impact of test use on longer-term patient outcomes such as morbidity and health-related quality of life (HRQoL), the EAG suggested that evidence linkage is needed. The following mechanism was used in identified cost-effectiveness models:

- Estimating the extent that cancer is detected at an earlier stage by EarlyCDT Lung (stage shift) based on the expected time to diagnosis (compared with standard care) and likelihood or time to preclinical stage progression, and
- Estimating long-term outcomes based on stage of cancer at diagnosis.

Full detail on the EAG's considerations for future modelling of the longer-term impact from increased or earlier detection of lung cancer can be found from page 138 in the diagnostics assessment report.

The size of any benefit of earlier or improved diagnosis by EarlyCDT Lung is reliant on the sensitivity of standard care tests (CT surveillance or PET-CT). The EAG commented that further evidence is needed to define the likelihood of malignant nodules being missed by the BTS surveillance schedule. The EAG also noted that there is no evidence that stage progression happens within the timeframe of CT surveillance. Evidence on the likelihood of stage progression with CT surveillance is therefore required to support a future assessment of EarlyCDT Lung.

The EAG also highlighted that earlier detection within the same disease stage could also be associated with long-term benefits, noting that this as an area in which further evidence is required. It also commented that the use of more disaggregated levels of staging categories (for example, stage IA1 distinct from stage IA2), could allow stage-shift based evidence linkage approaches to capture additional benefits that are currently not captured, reducing the impact of ignoring potential within-stage benefits.

Further data on the speed of preclinical stage progression in nodules was highlighted as potentially valuable by the EAG. This could include factors that could account for heterogeneity in stage progression, such as histology or other patient and nodule characteristics. An influence diagram reflecting the potential impact of heterogeneity can be found in the diagnostics assessment report on page 144.

The EAG highlighted that linkage to health outcomes from stage of diagnosis requires evidence on survival, health-related quality of life and costs conditional on disease stage at diagnosis. The EAG identified UK-specific evidence on these components. It commented that 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.

4 Ongoing studies

The EAG identified 2 potentially relevant ongoing studies of EarlyCDT Lung: 1 aiming to recruit 1,000 patients in China and the other screening US patients at high risk of developing lung cancer. It is unclear where in the pathway EarlyCDT is positioned in these trials, and whether it will be used to update risk scores as indicated. Therefore, the applicability of these trial results to NHS clinical practice is uncertain.

5 Summary

Clinical effectiveness

A meta-analysis of the EarlyCDT Lung diagnostic accuracy studies done by the EAG suggests that EarlyCDT Lung has 20.2% sensitivity (95% confidence interval (CI) 10.5 to 35.5) and specificity of 92.2% (95% CI 86.2 to 95.8). The EAG note that sensitivity is much lower than the estimate proposed by the company (41.3% sensitivity and 90.6% specificity from Healy et al. 2017). The EAG also note that the meta-analysis is limited by lack of data, potential for risk of bias and poor generalisability.

The EAG identified no evidence on the clinical impact of using the EarlyCDT Lung test, including how patients might be reclassified in terms of risk, or changes in their clinical management. To investigate the potential impact of test results on decision making, the EAG did a simulation study. This suggested that the EarlyCDT Lung test is unlikely to offer meaningful clinical improvement for low-risk nodules (0% to 10%). However, the EarlyCDT Lung may have some use in identifying malignant nodules among those classified as intermediate risk after Herder risk assessment. The EAG cautioned that these conclusions are from a simulation study, requiring strong modelling

assumptions. Use of the EAG's preferred accuracy estimates for EarlyCDT Lung meant a far lower impact of the test result on post-test risk estimates of malignancy, than if the company's preferred accuracy estimates were used.

The EAG identified no evidence to support the claim that the use of EarlyCDT Lung testing will result in the diagnosis of lung cancer at an earlier stage or improved patient outcomes compared to current practice.

Cost effectiveness

The EAG identified 2 cost-effectiveness studies on EarlyCDT Lung, but neither were considered relevant to the scope of the current decision problem. This was mainly because the test was modelled to identify whether a patient has a malignant nodule or not, rather than to upgrade a malignancy risk and that the diagnostic accuracy evidence used to inform the models may be overestimated (see page 19).

The EAG identified that the detection of malignant nodules at earlier stages of disease (stage shift) is a likely key driver of value of EarlyCDT Lung testing.

The extent of which the test adds information to currently used risk assessment tools, the impact of this on decisions about care, and how these decisions affect health-related quality of life and costs should be defined for future cost-effectiveness analyses.

Tables 5 and 6 summarise the EAG's suggested evidence requirements and modelling considerations to support a future assessment of the value of EarlyCDT Lung.

Table 5 Summary of clinical evidence requirements and considerations for modelling for a future NICE assessment of EarlyCDT Lung (adapted from table 28, page 149 of the diagnostics assessment report)

Key economic information	Evidence requirements	Considerations for modelling and evidence linkage

Population and subpopulations	• Reporting of detailed descriptions of the populations or subpopulations of interest • Better understanding of factors of heterogeneity of the disease prevalence, such as nodule characteristics (size and location), that are linked to likelihood of detection/magnitude of stage shift and outcomes.	• Explicit modelling of subgroups to reflect the different proposed positionings of the technology. • Relevant sources of heterogeneity related to malignancy risk, speed of nodule growth, speed of preclinical progression and long-term health outcomes, such as histological subtype.
Care pathway	• Reporting of the flow of patients in the care pathway with consequential clinical actions to assess clinical impact. • Better understanding of how risk of malignancy, and other factors, impact management decisions.	• Variation in management decisions and how this is related to risk of malignancy. • Variation from personalisation of care (that is, 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 populations, subpopulations or groups, demonstrating consistent benefit including in accuracy and in the validity of post-test risk scores. • Potential sources of heterogeneity should be examined. • 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 comparison to relevant NHS practice) 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 comparison to relevant NHS practice) 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).	• 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, such as the potential for increased detection.	• Explore the plausibility and relevance of including other value components in analyses.

Table 6 Summary of cost-effectiveness considerations for modelling for a future NICE assessment of EarlyCDT Lung (adapted from table 28, page 149 of the diagnostics assessment report)

Key economic information	**Key considerations**
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	• HRQoL 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.

Abbreviation: HRQoL, health-related quality of life.

6 Issues for consideration

Clinical effectiveness

The EAG's and company's preferred estimates for test accuracy differed, particularly for sensitivity. The EAG concluded that diagnostic accuracy of EarlyCDT Lung is uncertain and potentially at high risk of bias. It is unclear if the results of the ongoing studies will provide further test-accuracy data relevant for this assessment.

The EAG raised a concern that the proposed risk model for EarlyCDT Lung (Healey et al. 2017) may be based on biased estimates of accuracy. It commented that the risk model requires validation in independent cohorts. If its estimated risks do not match observed risks, a new model will be required, based on robust diagnostic accuracy data from new cohorts.

There is limited evidence on comparator tests used in the current pathway. In particular, the diagnostic accuracy of the Brock and Herder risk models is uncertain. Consequently, there is uncertainty about the diagnostic accuracy when combining the EarlyCDT Lung result with these risk models.

There are no data showing the impact of EarlyCDT Lung use on management decisions. Particularly in the 10% to 70% risk group, there is uncertainty about how risk of malignancy (used with other factors such as patient preference) determines choice of treatment or further diagnostic procedures within NHS clinical practice. The extent to which updating this risk based on the EarlyCDT Lung score would change these decisions is also uncertain. Based on

exploratory analysis, the EAG identified that the test score may only change care in people with low or intermediate-risk nodules, which are eligible for biopsy, and for people with nodules with a higher pre-test risk score intermediate regardless of biopsy feasibility.

Using the EAG's preferred accuracy estimates meant that use of the test had a far lower impact on updating risk of malignancy scores than if the company's preferred accuracy estimates were used.

Cost effectiveness

Greater characterisation of the population to be tested, including consideration of possible sources of heterogeneity is potentially important for future data collection.

Currently, the choice of treatment or further assessment based on risk of malignancy (with or without EarlyCDT Lung) in the intermediate-risk group is very uncertain, as described above.

The extent of overdiagnosis resulting from use of the test is unknown. That is, greater use of unnecessary invasive diagnostics or treatments on benign nodules and indolent nodules.

The main mechanism of value in existing cost-effectiveness studies of EarlyCDT Lung is of earlier detection, but there is no evidence that a potential earlier diagnosis with EarlyCDT Lung will result in earlier stage of cancer at diagnosis (compared with current care) and that improved patient outcomes will occur as a result of this.

The EAG highlight that there is currently limited evidence on stage progression during the CT surveillance period. Studies on the clinical consequences of CT surveillance (for example, the number of cancers identified and missed, delays in diagnosis and the possibility of tumour progression) are required for future EarlyCDT Lung assessments.

In the absence of direct data on the impact of the test on earlier detection of lung cancer (and consequent impact of this on longer-term outcomes), a linked evidence approach could be used. A future model would require data necessary to estimate the extent to which cancer is diagnosed at an earlier stage if the EarlyCDT Lung test is used, and the extent of long-term outcomes based on stage of cancer at diagnosis.

The EAG suggested the need for prospective cohort studies, with EarlyCDT Lung used in people with identified pulmonary nodules who are diagnosed and managed in line with the British Thoracic Society diagnostic pathway. This should include sufficient follow up to confirm malignancy by biopsy or surgery, or its absence, with at least 2 years' follow up without nodule growth. In addition to accuracy, clinical impact of the test could also be assessed, including impact of EarlyCDT Lung on risk classification, changes in clinical management, timing and tumour stage at detection and treatment of malignant nodules and use of CT or PET-CT scans. Further details are presented on page 157 in the diagnostics assessment report.

7 Equality considerations

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others.

People with lung cancer may be classified as having a disability and therefore protected under the Equality Act 2010 from the point of diagnosis. Over time, lung cancer rates in females have increased by almost a third, and rates in males have decreased by a third. There are also differences in the rates of lung cancer between ethnic groups. It is most common in White men and men of Bangladeshi family origin. Rates for men of Indian, Pakistani, Black Caribbean, Black African and Chinese family origin are lower. In women, lung cancer is more common in White women than in women from other ethnic groups. Sex and race are protected characteristics under the Equality Act

2010. The incidence and mortality of lung cancer are higher in deprived populations.

The EAG noted that studies have shown that the Brock model has inferior accuracy East Asian populations.

8 Implementation

EarlyCDT Lung testing is performed in a laboratory setting, some changes to laboratory infrastructure may be required to incorporate this testing pathway. As with all clinical laboratory tests, the development of local protocols and both internal and external quality assurance processes would be required.

9 Authors

Vera Unwin
Topic lead

Frances Nixon
Technical adviser

Thomas Walker
Technical adviser

10 Glossary

Standard care

The best available current clinical practice diagnostic test, against which the technology under assessment is compared.

Pulmonary nodule

A small, abnormal growth in the lungs that can be non-cancerous or cancerous.

Malignant

This is a cancer. The word malignant describes the fact that the cancer is not under normal control and that it has the potential to spread both locally and to distant areas.

Brock model

The Brock model is a multivariable model that estimates the risk that a pulmonary nodule on CT scan is lung cancer. By providing an estimate of nodule lung cancer risk, the Brock model can assist in determining appropriate follow up and management of pulmonary nodules detected on CT. The Brock model utilises key inputs including nodule size, location in the lung, age, smoking status, family history of lung cancer, sex, and nodule count.

Herder model

The Herder model predicts the risk of malignancy in solid pulmonary nodules using patient characteristics, nodules characteristics, and the degree of FDG uptake on PET-CT.

PET-CT scan

PET scanners work by detecting the radiation given off by a substance injected into the arm called a radiotracer as it collects in different parts of the body. In most PET scans a radiotracer called fluorodeoxyglucose (FDG) is

used, which is similar to naturally occurring glucose (a type of sugar). PET scans are combined with CT scans to produced detailed 3D images that highlight abnormalities in the body. They are used to confirm presence and spread of cancer.

Benign

This is a growth, which although abnormal in the body, is not a cancer. Although it may grow and spread locally it does not metastasise or spread to other areas of the body.

Incremental cost-effectiveness ratio (ICER)

The difference in the change in mean costs in the population of interest divided by the difference in the change in mean outcomes in the population of interest.

Indolent nodules

Nodules that are malignant but non-aggressive.

Staging

These are medical tests to establish the extent of a cancer. Staging is a way of describing the size and any spread of cancer and is an important factor in deciding the best treatment.