UK National Screening Committee

Final report on the cost-effectiveness of low dose computed tomography (LDCT) screening for lung cancer in high risk individuals

Version: 1.3 (Pre-External Quality Assessment)

Author: Exeter Test Group and Health Economics Group

Date: 30 November 2022

The UK National Screening Committee secretariat is hosted by The Office for Health Improvement & Disparities.

About the UK National Screening Committee (UK NSC)

The UK NSC advises ministers and the NHS in the 4 UK countries about all aspects of population screening and supports implementation of screening programmes. Conditions are reviewed against evidence review criteria according to the UK NSC's evidence review process.

Read a complete list of UK NSC recommendations.

UK NSC, OHID, Department of Health and Social Care, 39 Victoria Street, London, SW1H 0EU
www.gov.uk/uknsc

Published November 2022

Contents

List of tables

Plain English summary

The symptoms of lung cancer appear in the later stages of the disease. This is often when people are diagnosed and few treatment options are possible. Diagnosis at an earlier stage helps to achieve better results. If we can identify people with lung cancer before they develop symptoms, it is easier to treat the cancer.

Low dose computed tomography (LDCT) uses X-rays to show images through a person's body, and can show lung cancers. People have lower doses of radiation with LDCT than with a standard CT scan.

Many studies have looked at lung cancer screening for people who have smoked in the past or who are smokers. The studies aimed to find out if screening with LDCT would lead to better health outcomes compared to no screening.

The NHS looks at the value for money when considering new health interventions, such as treatments or tests. This involves comparing the benefits of the intervention with how much it costs. Computer models often help with working this out.

This document describes the updating of an existing model and presents the results. The updated version includes recent data from journals, expert opinions, and a completely new part of the model.

The updated model compares several screening approaches for lung cancer targeting people at high risk. It suggests that LDCT screening for lung cancer is very likely to be cost-effective for the NHS. Screening people for lung cancer using LDCT every year, or every two years, is more cost-effective than less intensive screening.

Executive summary

Purpose

This document describes the development and updating of an existing model-based economic evaluation of targeted low-dose CT scanning (LDCT) screening for lung cancer in the UK. In an interim report in March 2022 we reported the impact of updating the model parameters. This report incorporates major structural changes to the natural history model component. The cost-effectiveness results reported here supersede those in the interim report.

Background

A recently published review(1) identified 35 cost-effectiveness analyses of LDCT screening for lung cancer published since 2000. LDCT screening was generally found to be more effective and more costly than no screening. Reported incremental cost-effectiveness ratios (ICER) ranged from US$1464 to US$2 million per quality adjusted life-year (QALY) gained depending on policy question, setting, modeling approach, and evidence used. Four cost-effectiveness analyses CEA) based in the UK were identified – Whynes(2), Field(3), Hinde(4) and the original ENaBL (Exeter natural history-based economic model of lung cancer screening) report by Snowsill(5). All evaluated a single LDCT screen versus no screen. ICERs ranged from £8466 per QALY gained(3) to £28,169-£30,821 per QALY gained(5) depending on the eligible population. There was thus some uncertainty about the cost-effectiveness of LDCT screening for lung cancer.

During 2020-22 the UK National Institute of Health Research (NIHR) and the UK National Screening Committee (NSC) (via Public Health England) commissioned further development of a model-based economic evaluation of LDCT for lung cancer screening, ENaBL. An external review of the original model identified some targets for further development. An important concern was that the original natural history model estimated that at diagnosis there are more late stage cancers and fewer early stage cancers than is observed in data from trials and national statistics. This consequently meant that the potential benefits of lung cancer screening may not have been fully captured. Parameter estimates also needed to be brought up-to-date from 2016.

Focus

The objective of this report is to present the final cost-effectiveness results from ENaBL which incorporates both the updated natural history model component and the updated

parameter estimates relative to the original model developed in 2016. The intervention is targeted screening in high risk individuals for lung cancer by LDCT. Forty-eight screening strategies were compared to no screening and to each other. The strategies differed in terms of the frequency of screening (a single screen, three annual screens (triple screen), annual screening and biennial screening), and the populations who were eligible for screening (depending on age and predicted risk of lung cancer). We used a threshold of £20,000 per QALY to judge whether a strategy was cost-effective. We also used Incremental Net Monetary Benefit (INMB), valuing a QALY at £20,000 to present results. A NMB positive value above 0 indicates that a strategy is cost-effective in the same way that a cost per QALY being below £20,000 indicates this.

Recommendation under review

Targeted screening for lung cancer was recommended in June 2022 based on evidence including the interim report. This report is part of the further modelling work suggested to refine the recommendations.

Findings

Probabilistic analyses indicated that all screening strategies were associated with increased costs compared to a strategy of no screening, but were also associated with gains in QALYs. The incremental QALYs ranged from 0.0004 to 0.0132, with incremental costs ranging from £15 to £120 per person aged 55-80 years over a lifetime. Although the incremental QALYs and costs were very small, and would not generally be considered significant (a QALY gain of 0.0132 is less than 5 quality-adjusted days), the gains from screening are concentrated in those individuals who join the screening programme, and are diagnosed with lung cancer before it would have presented clinically. In this group the gains are clinically important.

In terms of cost-effectiveness, the single screening strategies do not look to be cost-effective at a willingness to pay of £20,000 per QALY gained. For the triple screening strategies, there is uncertainty as to whether they could be considered cost-effective. For the biennial and annual screening strategies, the results indicate that these strategies could be considered cost-effective compared to no screening.

Four annual LDCT screening strategies were identified to lie on the cost-effectiveness frontier. Thus, these strategies were estimated to give the maximum incremental net monetary benefit at a willingness to pay of £20,000 per QALY gained. For instance the INMB for an annual strategy inviting 55-75 year olds and screening those with a risk greater

than 1.5% (A-55-75-1.5%) was £142, [95% Credible Interval £302 to £14]. The cost per QALY was £8,517, [95% Credible Interval £4,119 to £18,287]

Recommendations on screening

The final ENaBL model confirms the interim report that targeted lung cancer screening with LDCT is cost-effective at a threshold of £20,000 per QALY.

The final ENaBL model adds detail in indicating that annual and biennial strategies are more cost-effective than less intensive screening strategies.

Evidence uncertainties

The modelling employed extensive sensitivity analyses, particularly probabilistic sensitivity analyses, to explore the impact of uncertainty. The findings were robust and we are thus certain that the cost-effectiveness will remain in the region considered cost-effective even taking uncertainty into account for annual and biennial strategies. The 95% credible intervals for INMBs on the annual screening strategies on the cost-effectiveness frontier do not include 0.

Further modelling will be required to contribute information on feasibility and budget impact.

Further modelling will be required to assess the impact on cost-effectiveness of adjuncts to targeted lung cancer screening with LDCT such as incorporation of smoking cessation or systematic treatment of incidental findings.

A further developed model could be useful for making initial assessments of the impact of possible future modifications of a targeted lung cancer screening programme.

Acknowledgements

We gratefully acknowledge the inputs from the clinical and methodological working groups who provided feed back on the original ENaBL model. We particularly thank David Baldwin for his inputs throughout the model development process. These acknowledgments do not necessarily indicate that those involved agree with the conclusions of the report.

Funding

The work reported was jointly funded by the National Screening Committee and the UK National Institute of Health Research.

Conflicts of interest

None

Introduction and approach

Purpose

The introduction which follows concentrates on the history of the project which this document reports – the further development and up-dating of the Exeter natural history-based economic model of lung cancer screening (ENaBL). For general background on the nature of lung cancer, the rationale for screening for it and important principles concerning the evaluation of new screening programmes we refer readers back to our original HTA report in 2018(5).

Background

Previous research on cost-effectiveness

A recently published review(1) identified 35 cost-effectiveness analyses of LDCT screening for lung cancer published since 2000. LDCT screening was generally found to be more effective and more costly than no screening. Reported ICERs ranged from US$1464 to US$2 million per QALY gained depending on policy question, setting, modeling approach, and evidence used (see Table 1). Four CEAs based in the UK were identified – Whynes(2), Field(3), Hinde(4) and the original ENaBL report by Snowsill(5). All evaluated a single LDCT screen versus no screen. ICERs ranged from £8466 per QALY gained(3) to £28,169-£30,821 per QALY gained(5) depending on the eligible population.

Snowsill also evaluated annual screens for 3 years (referred to as triple screen), and annual and biennial screens for given age ranges. For the triple screen only one strategy was on the cost-effectiveness frontier (ICER vs no screening: £40,034/QALY)(5). None of the annual and biennial strategies, were estimated to be on the cost-effectiveness frontier. An interim update analysis was conducted in which many of the parameters of ENaBL were updated, but the natural history model (which determines the risks of lung cancer and how these are affected by screening) was not changed from the original report. This interim update analysis produces ICERs more favourable to LDCT screening ranging from £1,529 per QALY (for 3 annual screens) to £4,385 per QALY (for annual screens between 55 and 80 years old), with no biennial screening strategies on the cost-effectiveness frontier. Note that ENaBL is the only UK-based model that has evaluated annual and biennial LDCT screening strategies. The cost-effectiveness analysis in the original ENaBL report (and by extension the interim update) has been criticised for relying on a natural history model which produces results out of line with observations in trials and routine clinical practice and this is indeed a limitation of those analyses.

Peters(1)reported on the variability of modelling approach, and concluded that those models incorporating a natural history component for lung cancer were more likely to adequately address critical appraisal items, but stressed that these are difficult to validate appropriately. The modelling approach taken by Whynes(2), Field(3) and Hinde(4)are similar, essentially using a decision tree approach where the effectiveness of LDCT screening is represented by an explicit stage shift at diagnosis – with those diagnosed via screening assumed to be diagnosed at an earlier stage. Consideration of overdiagnosis in these analyses is limited, pre-determined estimates of lead-time are assumed and few sensitivity analyses are reported. Although the data informing Whynes(2) is hypothetical, due to a lack of trial data at that time, the data used in Field(3) and Hinde(4) are from the UK. Snowsill(5) use a DES model incorporating a natural history model. This approach has advantages of implicitly considering issues such as lead-time bias and overdiagnosis, and is calibrated on data from the largest lung cancer screening trial conducted so far (NLST).

Among the other published CEAs, Peters identified a number of studies using models with a natural history component. Consideration of these models may provide context for the ENaBL model. Four of these are part of the US National Institute for Health Cancer Intervention and Surveillance Modeling Consortium and as such are registered with the Cancer Intervention and Surveillance Modeling Network (CISNET). They are MISCAN-Lung, lung cancer policy model (LCPM), lung cancer outcomes simulator (LCOS) and the model from the University of Michigan, see Table 1.

The MISCAN-Lung model has been used to evaluate LDCT screening versus no screening in Canada(6) and Switzerland(7). In Canada, annual screening strategies were associated with ICERs of Can$39,000/LY to Can$64,500/LY (cost year 2015) depending on age and smoking history of eligible population(6). In Switzerland, evaluation of annual screening strategies lead to ICERs of €30,500/LY to €48,500/LY depending on age and smoking history of eligible population were estimated(7) Comparison with ENaBL is difficult as analyses per QALY gained were not reported. However, ICERs per QALY gained would likely be greater than those reported per LY gained as the main effect would be to apply population norms (i.e., less than perfect health-related quality of life on average) to life expectancy gains. It is worth noting that, although not stated explicitly, the WTP per QALY gained in Canada is generally thought to be around Can$50,000.

The LCPM was used in McMahon(8) to evaluate single and annual screening vs no screening in the US with a cost-year of 2006. Depending on gender, age group and smoking history of eligible population, ICERs (compared to no screening) ranged from US$144,000 - $207,000/QALY for single LDCT screens and $110,000/QALY - $203,000/QALY for annual screening. Criss(9) recently evaluated annual screening in the

US starting at age 55 years old, with different upper age limits using all four of these CISNET models (cost year 2018). They reported average ICERs across the 4 models of $49,200/QALY (stopping at age 74), $68,600/QALY (stopping at age 77), $96,700/QALY (stopping at age 80), see Table 1. As with Canada, it is generally thought that the WTP per QALY gained in the US is around US$50,000. However, the USPSTF does not consider cost-effectiveness in their decision-making.

As noted in Peters(1), making comparison between different evaluations of cost-effectiveness is not straightforward, due to the multiple sources of heterogeneity. However, simple, naïve, comparison of ICERs between McMahon and Criss suggest that LDCT is seemingly more cost-effective in the US now than it was 15 years ago.

The UK-based studies report lower ICERs than those based on the CISNET models. There are many differences between the analyses, including the modelling approach and the strategies evaluated. With the exception of Snowsill(5), the UK-based studies use decision tree approaches, while the CISNET models use more complex models incorporating the natural history of lung cancer. The UK-based studies have focussed on single screening, while the CISNET models have focussed more on the cost-effectiveness of annual screening. McMahon(8) and Snowsill being the exceptions. The original ENaBL model assessed both, finding that a single screen was generally more cost-effective than annual screening.

As this summary of UK-based and CISNET models suggest, analyses based on more complex natural history-based models have tended to produce higher ICERs than those using decision tree approaches. Peters(1) also found that they tended to address more of the challenges of evaluating cancer screening programmes than less complex models.

Table 1 Characteristics and results of UK-based models and CISNET registered models by screening frequency

Model	Country	Cost-year	Eligible population	Incremental costs (vs no screening) per person	Incremental effects (vs no screening) per person	ICERs (vs no screening)
Single screen						
Original ENaBL(5)	UK	2016	Aged 60-75 years, ≥3%	£23	QALYs: 0.0008	£28,169 per QALY
			Aged 55-75 years, ≥3%	£3	QALYs: 0.0001	£28,784 per QALY
			Aged 55-80 years, ≥3%	£6	QALYs: 0.0001	£30,821 per QALY
Whynes(2)	UK	2004	Men aged 61 years at high risk	£201	QALYs: 0.01	£14,000 per QALY
Field(3)	UK	2016	Adults aged 50–75 years, at =>5% risk of lung cancer	£565,498*	QALYs: 66.8*	£8466 per QALY
Hinde(4)	Manchester	2015	55-74yrs ever smokers with 6-year lung cancer risk of ≥1.51%	£40	QALYs: 0.004	£10,069 per QALY
LCPM McMahon(8)	US	2006	Aged 50-70, 60-74, 70-74 with current & former =>20 pack-year history	US$1,778 to US$3,637	QALYs: 0.009 to 0.022	US$144,000 to $207,000/QALY
Annual screening (for 3 years)						
Original ENaBL(5)	UK	2016	Aged 55-80 years, ≥3%	£17	QALYs: 0.0002	£40,034 per QALY
Annual screening (for age group)						

Original ENaBL(5)	UK	2016	Various			None on the efficient frontier
MISCAN(6)	Canada	2015	10-40 pack-years. 10 - 20 years since smoking cessation.	Can$498 to $2067	LYs: 0.013 to 0.032	Can$39,000 to $64,500 per LY
MISCAN(7)	Switzerland	2015	10-40 pack-years. 10-20 years since smoking cessation.	€641 to €1885	LYs: 0.021 to 0.039	€30,500 to €48,500 per LY
LCPM McMahon(10)	US	2006	Aged 50-70, 60-74, 70-74 with current & former =>20 pack-year history	NR	NR	$110,000/QALY $203,000/QALY depending on gender, age group and smoking history of eligible population.
LCOS Toumazis(11)	US	2019	20-40 pack-years.10-20 years smoking cessation	US$903 to $2,391	QALYs: 0.0161 to 0.0193	US$55,968/QALY to US$124,147/QALY depending on age and smoking history of eligible population and whether disutility for indeterminate results included
Criss(9)	US	2018	Lower age limit 55 years.	US$870 to $980	QALYs: 0.019 to 0.021	Average across the 4 models: $49,200/QALY (stop at age 74), $68,600/QALY (stop at age 77), $96,700/QALY (stop at age 80)
Biennial screening						
Original ENaBL(5)	UK	2016	Various			None on the efficient frontier
MISCAN(7)	Switzerland	2015	30-40 pack-years.	€324 to €6100	LYs: 0.013 to 0.020	€25,500 to €31,000 per LY
LCOS Toumazis(11)	US	2019	30-40 pack-years, 10-15 years	US$282 to $1,033	QALYs: 0.0065 to 0.0134	US$43118/QALY – US$76909/QALY depending on age and smoking history of eligible

			smoking cessation.			population, and inclusion of disutility for indeterminate results
MISCAN(6)	Canada	2015	Various			None on the efficient frontier
Triennial screening						
MISCAN(7)	Switzerland	2015	30-40 pack-years.	€333	LYs: 0.012	€27,374 per LY

*For the total population, not per person;

LCOS, Lung Cancer Outcomes Simulator; LCPM, Lung Cancer Policy Model; LY, life-year; QALY, quality-adjusted life-year

History of the project

Original ENaBL model

In November 2016 researchers at the University of Exeter were commissioned by the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) programme to investigate the effectiveness and cost-effectiveness of low-dose computed tomography (LDCT) for lung cancer screening (HTA 14/151/07). The project was completed in summer 2017 and published December 2018(5). An independent model-based economic evaluation was undertaken, resulting in ENaBL. It consists of a discrete event simulation (DES) model incorporating a natural history model for lung cancer. Four LDCT screening frequencies were evaluated, in addition to no screening, in 12 different populations defined by age range, and predicted risk of lung cancer. The findings of the preliminary model indicated that a single (one-off) LDCT screen could be considered cost-effective at conventional willingness to pay thresholds, but that there was “significant uncertainty about the effects of costs and the magnitude of benefits”(5).

The NIHR and the UK National Screening Committee (NSC) (via Public Health England) commissioned further development of the health technology assessment of LDCT for lung cancer screening to reduce the uncertainty by incorporating additional evidence up-dating parameters and by addressing concerns surrounding the model developed for the cost-effectiveness analysis.

ENaBL used within it a model of the preclinical development of lung cancer so that screening programmes can be simulated which have not been evaluated in clinical studies. This natural history model incorporates the risks of developing preclinical (occult) lung

cancer, progression of preclinical lung cancer (through seven lung cancer stages based on American Joint Committee on Cancer (AJCC) and Union for International Cancer Control (UICC) 7th edition; IA, IB, IIA, IIB, IIIA, IIIB and IV), and the presentation (symptomatic or incidental) of lung cancer. External validation of ENaBL showed that its natural history component resulted in stage distributions in the presence or absence of screening that are not well matched to the stage distributions observed in LDCT trials and national statistics, with an overestimation of late stage cancers and underestimation of early stage cancers. This consequently meant that lung cancer survival may have been underestimated, and that the potential benefits of lung cancer screening may not have been captured.

Process for updating ENaBL

Towards the end of 2020, a number of stakeholder meetings were held between the University of Exeter team, members of the NSC, and clinical and modelling experts to discuss the ENaBL model and prioritise the work involved to update and further develop the model (referred to as the Modelling Task and Finish Group meetings).
Resulting from these consultation meetings, the University of Exeter team proposed to:

1. Re-consider, and update where appropriate, all parameters in the original model.
2. Create, calibrate and validate a new natural history model to underpin ENaBL.

Based on feedback from the Modelling Task and Finish Group meetings, a number of assumptions to revisit in the natural history model were highlighted and are shown in Table 2. The aim was to calibrate the natural history model with and without these revised assumptions.

Table 2 Potential revisions to assumptions in the natural history model

ID	Assumption in original ENaBL(5)	Likely alternative assumption	Expected effect of revising assumption
A	No lung cancer mortality without prior diagnosis of lung cancer	A certain proportion of people with preclinical lung cancer in Stage III or IV will die from lung cancer without first being diagnosed	Unclear
B	Equal sensitivity of LDCT screening across preclinical lung cancer stages	Separate sensitivity estimates for preclinical Stage IA versus preclinical Stages IB–IV	Sensitivity may be increased for late stages and lowered for earlier stages; Effect on cost-effectiveness will depend on what

			proportion of early stage cancers are detected if sensitivity is lowered, and how many early stage cancers would be "overdiagnosed".
C	The hazard function for preclinical disease incidence is dependent only on age and sex [a]	The hazard function for preclinical disease incidence will additionally incorporate information about smoking history [b]	Will marginally benefit annual and biennial screening [c]
D	UICC/AJCC 7th edition staging	UICC/AJCC 8th edition staging, including separating Stage IA into Stage IA1, IA2 and IA3 and merging Stages IIA and IIB into Stage II, Stages IIIA and IIIB into Stage III. Stages IIB and IV to remain as they were.	Will likely benefit screening programmes

AJCC, American Joint Committee on Cancer; ENaBL, Exeter natural history-based economic model of lung cancer screening; UICC, Union for International Cancer Control.

Notes: [a] The existing model specifically assumes that once patients have been selected for having ever smoked (i.e., current or former smokers) they face the same risk of developing preclinical lung cancer for each year of life after controlling for age and sex. This does not mean that all patients will get preclinical lung cancer at the same age – event times in the model are sampled from probability distributions.

[b] This could be, e.g., pack years, comprehensive smoking index (CSI), linear prediction from LLPv2 or other risk prediction tool.

[c] The existing model reverse-engineers a distribution for the LLPv2 risk scores from the modelled lung cancer outcomes within 3 years of the model start, which has the effect of meaning those with higher risk scores are more likely to develop lung cancer within 3 years. However, after 3 years has passed, the probability of a simulated patient developing lung cancer is not affected by their simulated risk score. Revising this assumption would mean that those with higher baseline risk could continue to have elevated lung cancer risk, which would benefit annual and biennial screening programmes in the simulation.

At the beginning of 2021, updating of ENaBL began. During this process, further clinical input was sought, particularly in the re-development of the natural history model. This

consisted of multiple meetings and correspondence with clinical experts. As part of this, the appropriateness of additional assumptions in the natural history model were raised. This included heterogeneity associated with lung cancer. This can be partly explained by histology, but also, with the progression and presentation of lung cancers to reflect indolent and fast-growing tumours.

Not accounting for indolent tumours, it is assumed that all cancers identified will impact clinically on a patient, and require some intervention. This may not be the case for some very slow growing cancers where the individual would not have experienced any impacts during their lifetime. Not accounting for these slow-growing cancers could lead to overestimation of the (cost-) effectiveness of LDCT screening compared to no screening. This is because not every cancer identified would have impacted on the individual, as they may die from other causes before the lung cancer has any clinical impact (overdiagnosis). Thus, any intervention would be unnecessary, and incur unnecessary costs and potential impacts on quality of life for the individual.

By not accounting for fast-growing tumours, the model will not adequately estimate the number of interval cancers diagnosed in a screening programme, i.e. cancers diagnosed between screens, that will not have been present at the time of screening. If these cancers are not modelled appropriately, the proportion of screen-detected cancers will likely overestimated, leading to overestimates of the (cost-) effectiveness of LDCT screening compared to no screening.

To account for heterogeneity in the updated natural history model, non-small cell lung cancer (NSCLC, stages I-IV), and small cell lung cancer (SCLC, limited and extensive) were to be modelled separately, with additional heterogeneity parameters for NSCLC.

Interim report March 2022

Due to the number of assumptions to be revisited, a completely new natural history model was developed. However, because of the complexity of this, we were delayed in our delivery of the updated model. Thus, at the request of the UK NSC, we prepared an interim report, looking at just the effect of updating parameters within the original model. Reporting interim, or emerging, findings in this way was felt to be a timely contribution to discussion of the UK NSC recommendation on lung cancer screening in mid-2022. There was however no updating of the natural history model used in the original report, therefore all criticisms and limitations of this part of the model remained. These have been fully addressed in the version of the model reported in this report.

The findings of the interim report were that updates to parameter values and limited revisions to the structure of the DES model led to 4 LDCT screening strategies lying on the cost-effectiveness frontier in base case analyses. Screening strategies were estimated to be more effective than no screening, suggesting a QALY gain of 0.006 to 0.0029 per person, depending on the strategy. Although such gains would not generally be considered significant, these gains are concentrated in people who join the screening programme (ranging from 3.6% to 12.6% of the population), are diagnosed with lung cancer at an earlier stage and, therefore, receive more substantial health benefits. Screening strategies were estimated to be more costly than no screening, with an additional £16 to £126 cost per person. The results from this interim update suggested that LDCT screening would be cost-effective compared to no screening at a willingness to pay of £20,000 per QALY gained. For the strategies on the cost-effectiveness frontier, the model estimated that LDCT screening reduced lung cancer mortality by 3.1% to 5.3% compared to no screening. There was a pattern of increasing cost and QALYs as the number of screens in the programme design increases. The same pattern was observed in respect to lung cancer risk: lowering the threshold leads to increasing costs and QALYs.

Objectives

The objective of this report is to present model results using an updated natural history model and updated parameter estimates, where relevant, as applied to the original ENaBL model(5). The clinical effectiveness of LDCT screening is not addressed in this report, and readers are referred to the Rapid Review up-date commissioned by the UK NSC secretariat. This report only deals with aspects of effectiveness which impact on the estimation of cost-effectiveness.

Model overview

Decision problem

The decision problem assessed is the same as in the original report(5): to evaluate the cost-effectiveness of 48 LDCT screening strategies (defined by screening frequency and eligibility criteria) and a strategy of no LDCT screening, in a population at high-risk of lung cancer in the UK.

Population

Those eligible for LDCT screening were assumed to be individuals aged 55-80 years with a history of smoking (current or former). Only those with a risk of lung cancer above a specified threshold as calculated by the Prostate Lung Colorectal Ovarian lung cancer risk prediction model ($PLCO_{m2012}$)(12) (1.5%, 2.5% or 5%), were invited for screening. It was further assumed that only those individuals with performance score (PS) 0-2 would take-up the offer of screening. This is based on data from Crosbie 2019(13) where only 1.5% of participants accepting LDCT screening in the Manchester Lung Health Check (LHC) pilot had a PS >2.

Setting and location

As in the original report, the evaluation is based in the NHS in the UK.

Screening programmes

No changes were made to the decision problem evaluated in the original report(5). Four screening strategies were modelled, and compared to a strategy of no LDCT screening:

- A single, one-off LDCT screen (as in the protocol for the UK Lung Cancer Screening (UKLS) trial(3))
- Triple LDCT screening – 3 consecutive annual screens (as in the protocol for the NLST(14))
- Annual LDCT screening (as recommended by the US Preventative Services Task Force (USPSTF)(15))
- Biennial LDCT screening

In addition to screening frequency, strategies were evaluated assuming screening was offered at different lung cancer risk thresholds (as described above), and different age ranges for individuals. The lower age limits were assumed to be 55 or 60 years old, with

upper age limits of 75 or 80 years old. Thus, there were 48 distinct LDCT screening strategies evaluated and compared with no screening.

General approach

As in the original report(5), a microsimulation model was developed, using a discrete event simulation approach to sample individuals with a range of baseline characteristics. These individuals were concurrently simulated across different screening strategies, defined by screening frequency, as well as a no screening strategy (which represented current practice). The screening strategies were further defined by the population eligible to join each strategy, in terms of age and predicted risk of lung cancer. This resulted in 48 different screening strategies to be evaluated and compared to no screening.

However, to model the impact of screening, estimates of what would happen to an individual if they received screening and what would happen if they did not receive screening are needed. To do this appropriately, knowledge of the natural history of lung cancer for that individual is required. Observable data from trials and national statistics/registries, only provide information from the point of diagnosis, whether through screening or through clinical presentation.
To effectively model the impact of screening, need to know:

- Whether individual has cancer at the start of the screening programme
- The probably of developing cancer throughout the screening programme
- How quickly pre-clinical cancer will progress
- The probably that pre-clinical cancer will be identified through screening
- The probably that pre-clinical cancer will be identified clinically (in the absence of screening)
- The probability of dying from undiagnosed lung cancer (very rare)

In the original report(5), a model of the natural history of lung cancer was developed to estimate the risk of these unobservable events. It was calibrated to data from the NLST. However, due to a number of limitations with the natural history model, including distributions of stage at diagnosis not being consistent with empirical data, a completely new natural history model has been built. Details of the new natural history model are given in the next Chapter, “Natural history model”. The outputs from the natural history model include estimates of the sensitivity of LDCT by stage at diagnosis, the prevalence, incidence and progression of lung cancer. These outputs are then used as parameter

inputs to the discrete event simulation model described in Chapter “Economic evaluation of targeted lung cancer screening”.

Natural history model

Natural history modelling is the preferred approach for model-based cost-utility analyses of screening programmes(16). Natural history modelling allows economic evaluations to include options that have not been evaluated in primary research, including considering alternative eligibility criteria (e.g., age and predicted cancer risk), different screening schedules and even alternative screening technology with some assumptions. We have developed a *de novo* natural history model for this health technology assessment of lung cancer screening. This model has a number of key components:

- Prevalence of lung cancer at baseline
- Preclinical incidence
- Preclinical disease progression
- Screening
- Clinical presentation
- Death from other causes
- Death from undiagnosed lung cancer.

Each of these components is described in greater detail in subsequent sections. The model incorporates smoking history as a risk factor for prevalence, preclinical incidence and death from other causes. The natural history model described in this chapter does not include post-diagnosis survival of lung cancer, which is outside the scope of the data which is modelled. Post-diagnosis survival is included in the economic model based on observed survival data.

The key data source for the natural history model is patient-level data from the US National Lung Screening Trial (NLST)(17). NLST was a randomised controlled trial of low-dose computed tomography screening for lung cancer versus screening by chest X-ray; it included a baseline (prevalent) screen and two follow-up (incident) screens, spaced apart by 12 months. These repeated (imperfect) observations of the state of the lungs of the participants, plus several years of post-screening follow-up, give us the opportunity to model the prevalence, incidence, presentation and detection of lung cancer.

We have adopted a fully Bayesian approach, which requires the specification of a prior distribution of the model parameters (representing any knowledge or intuition about which parameter values are more or less credible before observing the data) and a likelihood function, which is the probability of the observed data given a particular set of parameter values. Once these are specified, numerical Bayesian methods are applied to determine the posterior distribution of the parameters, i.e., the appropriate distribution of the parameters given the specified prior distribution and data.

Most existing natural history models have used aggregate level data for calibration, even if they have used stochastic individual patient simulation to estimate the likelihood function(18-20). In contrast, we use individual-level data and deterministic individual patient simulation in our calculation of the likelihood function.

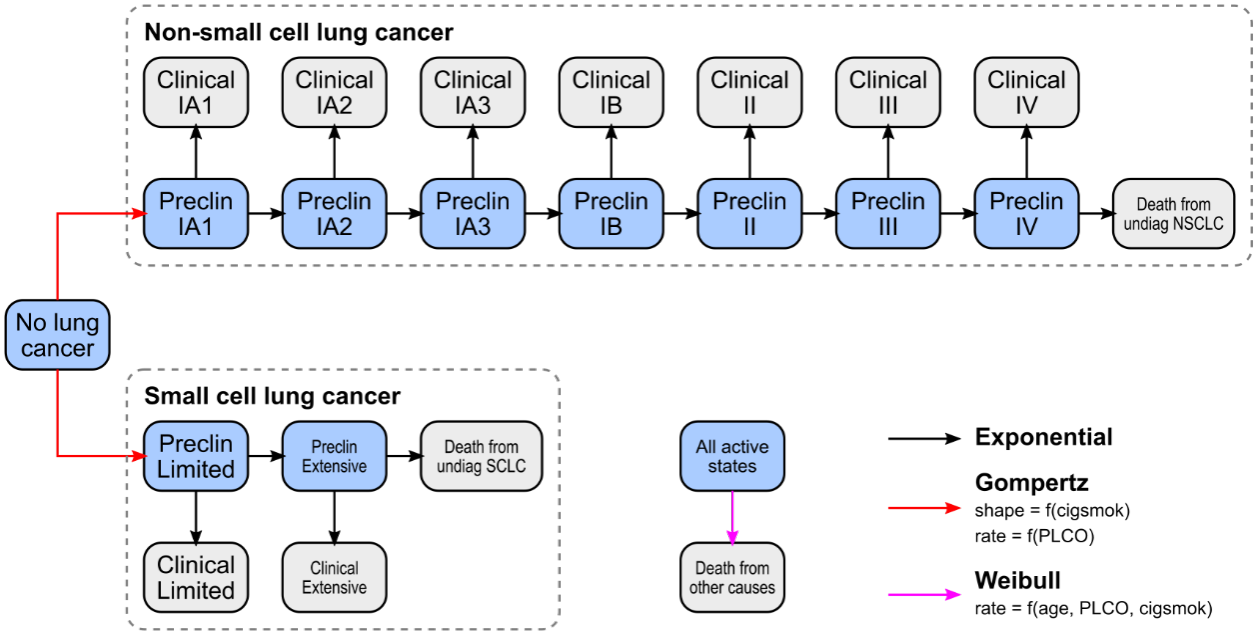
We have used the Stan No-U-Turn Sampler for our analyses(21). In this chapter we describe the components of the natural history model in detail, give further details of the data, describe the prior distribution for the model parameters, describe the calculation of the likelihood function, describe the procedure for calibration and validation, and show the results of the calibration and validation.

Model components

Lung cancer states

We divide lung cancer into small-cell lung cancer (SCLC) and non-small-cell lung cancer (NSCLC). As a simplification we include all lung cancers which are not SCLC in NSCLC. For NSCLC we include seven cancer stages according to the 8th edition staging: IA1, IA2, IA3, IB, II (i.e., IIA/IIB), III (i.e., IIIA/IIIB/IIIC) and IV (i.e., IVA/IVB). It is assumed that cancer progresses through these stages in sequence. These stages were chosen since the objective of lung cancer screening is to identify cancer in the earliest stages to maximise survival. We did not include further stages because this could make the assumption of sequential progression unrealistic. For SCLC we included only limited and extensive stages, since the objective of lung cancer screening is generally not to identify SCLC which is uncommon and generally very aggressive. Figure 1 shows a schematic of the natural history model, with the NSCLC and SCLC health states.

Figure 1 Schematic of the natural history model



Smoking history

Throughout the model, smoking history (and other lung cancer risk factors) is encapsulated in a single number, which is the linear predictor from the $PLCO_{m2012}$ risk prediction tool(12). The $PLCO_{m2012}$ risk prediction tool was adopted because it has been shown to have good discriminative performance(22, 23). We do not use $PLCO_{m2012}$ to directly predict the prevalence or incidence of lung cancer, it is only used as a convenient representation of smoking history and other lung cancer risk factors.

Prevalence

Our natural history model includes a prevalence component, i.e., a component which estimates the probability that an individual has an occult lung cancer (and characteristics of that lung cancer) at baseline, based on their smoking history.
If the individual's $PLCO_{m2012}$ linear predictor is x_i then the probability that the individual is in state j at baseline is given by a multinomial regression equation:

$$p_j(x_i) = \frac{\exp\left(\eta_j(x_i)\right)}{\sum_k \exp\left(\eta_k(x_i)\right)}$$

$$\eta_j(x_i) = \begin{cases} 0, & j = \text{No lung cancer} \\ \alpha_j + \beta_{NS} x_i, & j = \text{NSCLC Stage IA1, IA2, ..., IV} \\ \alpha_j + \beta_S x_i, & j = \text{Limited SCLC, Extensive SCLC} \end{cases}$$

There is a separate intercept for each stage of lung cancer, and there is a shared gradient term for the $\text{PLCO}_{\text{m2012}}$ for NSCLC and another shared gradient term for SCLC.

Preclinical incidence

In our natural history model, preclinical lung cancer incidence occurs when an individual goes from having no detectable lung cancer to having a lung cancer in the earliest stage, i.e., Stage IA1 NSCLC or Limited SCLC.

The $\text{PLCO}_{\text{m2012}}$ linear predictor is conventionally used to calculate a medium-term risk of a lung cancer diagnosis assuming a logistic regression. We assume instead that the linear predictor is proportional to the instantaneous log hazard rate of developing preclinical lung cancer. If the $\text{PLCO}_{\text{m2012}}$ predictor for an individual at age x is $\text{PLCO}_i(x)$ then (assuming no changes in smoking behaviour) their $\text{PLCO}_{\text{m2012}}$ predictor at age $x + \delta x$ is $\text{PLCO}_i(x) + \beta_f \delta x$ if they are a former smoker or $\text{PLCO}_i(x) + \beta_c \delta x$ if they are a current smoker (where β_f and β_c are known and are combinations of the $\text{PLCO}_{\text{m2012}}$ coefficients for age and smoking duration/time since smoking cessation). As log hazard is linear this implies a Gompertz distribution for the time of preclinical incidence with shape parameter $m\beta_f$ or $m\beta_c$ (depending on smoking status) and rate parameter $\exp\{c + m \cdot \text{PLCO}\}$ (m and c are parameters identified in the calibration).

$$\ln h_i(t) = c + m \cdot (\text{PLCO}_i + \beta t)$$

$$h_i(t) = \exp\{c + m \cdot \text{PLCO}_i\} \exp\{m\beta t\}$$

We in fact have two m parameters and two c parameters as there are two lung cancer histologies.

Preclinical disease progression

As previously stated, lung cancer is assumed to progress in sequence through stages. The time to each preclinical disease progression event is assumed to follow an exponential distribution, with each distribution having its own rate parameter (i.e., six progression rates for NSCLC and one for SCLC). Furthermore, since NSCLC is itself heterogeneous, with some cancers being highly aggressive and others more indolent, we include random effects at the patient level (i.e., at the cancer level, since the model accommodates at most one lung cancer per patient). If $\lambda_1, \ldots, \lambda_6$ are the "baseline" progression rates for NSCLC, then the progression rates for a particular patient are given by:

$$\lambda_j^{(i)} = \lambda_j \omega_i, \quad \omega_i \sim \mathrm{LogN}\left(-\frac{\sigma_{\mathrm{RE}}^2}{2}, \sigma_{\mathrm{RE}}\right) \tag{1}$$

Note that the distribution of ω_i is such that $\mathbb{E}\omega_i = 1$.

Screening

In the natural history model it is not necessary to include imperfect test specificity – this can be estimated empirically from screening studies. We do include the sensitivity of LDCT for detecting lung cancer in the natural history model, as well as the sensitivity of chest X-ray (CXR). We assume that LDCT and CXR are perfectly sensitive for Stage IV NSCLC and Extensive SCLC to avoid parameter identifiability issues and because this assumption is considered clinically valid. We assume that LDCT and CXR sensitivity increase as cancers progress, i.e., the sensitivity for Stage IA2 cannot be lower than the sensitivity for Stage IA1 and so on.

We make no assumptions about the relative sensitivity of LDCT versus CXR – any differences are purely identified from the data. We also include a final screen threshold effect for NLST, which is applied consistently across sensitivity values for all cancer stages as an odds ratio, with one odds ratio for LDCT and one for CXR.

Clinical presentation

The time to clinical presentation (due to symptoms or incidental findings) is given by a stage-dependent exponential distribution. We separate the rates of clinical presentation into stage-specific terms which must be non-decreasing as cancer becomes more advanced (primarily reflecting the possibility of symptomatic presentation) and context-specific terms which do not depend on the stage. The contexts are NLST mid-screening and NLST post-screening, and the corresponding terms represent the possibility that behaviour in NLST may change between the screening and post-screening periods.

Death from other causes

To examine the risk of death from other causes in NLST, we performed frequentist survival analysis for the event of death, censoring at the time of lung cancer diagnosis where applicable. We experimented with a range of parametric distributions and coefficients to include, and found that the optimal model (as judged by the Akaike Information Criterion) was a Weibull model with hazard ratios for age, the $\mathrm{PLCO}_{\mathrm{m2012}}$ linear predictor (mean centred), current smoking status and the interaction of $\mathrm{PLCO}_{\mathrm{m2012}}$ with current smoking

status. We incorporated the results of this survival analysis as data (i.e., not parameters to be estimated) in the Bayesian analysis.

Death from undiagnosed lung cancer

In our natural history model it is possible to die from undiagnosed lung cancer when it is metastatic (i.e., Stage IV NSCLC or Extensive SCLC), with an exponential time-to-event distribution. In our datasets we would count someone whose lung cancer was documented only on the death certificate as having clinical presentation of lung cancer rather than dying from undiagnosed lung cancer, so this risk solely refers to those who die from lung cancer but in whom lung cancer is never identified. It is not possible to directly observe this quantity, so we use a prior distribution on the proportion of metastatic lung cancers which are never diagnosed (see Prior distributions).

Summary

The model contains 52 parameters, of which 16 relate to diagnostic test performance, 11 relate to the prevalence of cancer at baseline, 11 relate to cancer presentation, 8 relate to cancer progression, 4 relate to cancer incidence, and 2 relate to death from undiagnosed cancer.

Data source – National Lung Screening Trial

A standard anonymised individual-level dataset from the NLST was obtained from the NLST investigators. Only the first lung cancer identified in a participant was included. All 53,452 participants were included in our analysis, and the first lung cancer in any participant was included (2,058 lung cancers included, 92 lung cancers excluded as not the first lung cancer in a participant).

Lung cancers were generally staged according to the 6th and 7th editions of the AJCC/UICC staging manual, whereas our natural history model was specified in terms of the 8th edition staging manual since there is significant survival differences within Stage IA in the 7th edition.

For 36 participants, we imputed a stage from the recorded 6th edition stage as insufficient data was present from 7th edition. For 18 participants, insufficient data was available for any staging, so we performed a single random imputation of the stage based on the survival time (restricted mean survival pseudo-observations were calculated with a time horizon of

seven years and this was used as the sole covariate for a multinomial logistic regression from which stage probabilities were estimated).

For all lung cancers, we imputed the stage under the 8th edition by using the Tumour Node Metastasis (TNM) data present, plus lesion size estimates. The results of this are shown in Table 3.

Table 3 Lung cancer stages imputed from NLST data for the natural history model

7th edition	AJCC/UICC 8th edition stage						
	IA1	IA2	IA3	IB	IIA / IIB	IIIA / IIIB / IIIC	IVA / IVB
IA	125	384	109				
IB				135	35		
IIA					129		
IIB					49	26	
IIIA						272	
IIIB						103	
IV							637
Not available	2	4	12	2	1	21	13

Prior distributions

Prior distributions are used to represent prior knowledge or intuition about the values a parameter may take. It is common in Bayesian analyses to use vague priors which do little more than ensure minimal probability mass is placed in areas of parameter space which are exceedingly unlikely to contain the parameter value. However, vague priors can result in failure to identify a posterior distribution in some cases (typically when data are not particularly informative for a parameter), so in these cases a weakly informative prior may be necessary. It may also be desirable for an informative prior to be used because good quality evidence exists from data which is not incorporated in the Bayesian analysis.

Sensitivity - Non-small cell lung cancer

We use the same prior distributions for low-dose CT screening and chest X-ray sensitivity to avoid introducing bias into estimates of the difference in screening performance between these technologies. A recent systematic review found three studies reporting the sensitivity of chest X-ray judged to be at low risk of bias(24). The studies generally found sensitivity close to 0.8. In the natural history model we include sensitivity parameters according to the stage of lung cancer, so priors were needed for each stage. We adopted prior distributions

with mean 0.8 for Stage II NSCLC, 0.9 for Stage III NSCLC and 0.7 for Stage I NSCLC (including all substages). The priors were defined to have an effective sample size of 20. If Beta distributions had been used, this would mean, for example a distribution of Beta(14, 6) to achieve a mean 0.7 and effective sample size 20, however logit-normal distributions were used, and so these were constructed from the moments (mean and variance) of the relevant Beta distributions. See Table 4.

Table 4 Prior distributions used for the NSCLC sensitivity parameters

Stage	Mean	α	β	μ	σ
I	0.7	14	6	0.894	0.496
II	0.8	16	4	1.476	0.559
III	0.9	18	2	2.393	0.716

Sensitivity – Small cell lung cancer

For limited small cell lung cancer we used a uniform prior on [0, 1], i.e., any value for sensitivity for SCLC was equally plausible under the prior distribution.

Sensitivity – Final screen threshold effect

We account for the possibility that investigators may have used a lower threshold in the final screen in NLST as it could be seen as the “last chance” to catch a lung cancer while operable. This takes the form of a non-negative adjustment on the logit-sensitivity scale which is common across all stages and histologies. We used a prior distribution of $\mathrm{LogN}(-2,1)$.

Preclinical incidence

The key parameters for preclinical incidence are c and m. m determines how important smoking history is to the future risk of preclinical lung cancer incidence, while c determines the risk for a “median” smoker (i.e., with median PLCO in the NLST study). These parameters interact substantially such that we consider a prior joint distribution is more appropriate. Recall that the hazard rate is given by

$$\ln h_i(t) = c + m \cdot (\mathrm{PLCO}_i + \beta t)$$
$$h_i(t) = \exp\{c + m \cdot \mathrm{PLCO}_i\} \exp\{m\beta t\}$$

NSCLC: We considered the hazard rate at 4 years from baseline in relation to the observed rate of lung cancers in NLST. Between 4 and 6 years from randomisation in the chest X-ray arm approximately 300 lung cancers were diagnosed, out of a population of approximately 26,000, suggesting a rate of approximately 6 per 1,000 person years. We next considered a grid of c and m values and calculate the resulting average hazard rates (averaged over the PLCO distributions of current and former smokers) at 4 years. If we count rates over 20 per 1,000 person years or below 0.5 per 1,000 person years as extremely unlikely, then a plausible prior distribution is one where m is between 0 and 4 and c is $-5-\frac{5m}{4}\pm 2$. We therefore apply a bivariate normal distribution for (m,c) with $\mu=(2,-7.5)^{\mathsf{T}}$ and

$$\Sigma=\begin{pmatrix}1 & -4/5\\ -4/5 & 25/16\end{pmatrix}.$$

This ensures approximately 95% of points have $m\in[0,4]$, and approximately 94% have $c\in\left[-7-\frac{5m}{4},-3-\frac{5m}{4}\right]$.

SCLC: We used a bivariate normal distribution for (m,c) with $\mu=(2,-10.5)^{\mathsf{T}}$ and Σ chosen the same as for NSCLC, i.e., the prior distributions were the same but with c shifted so its expected value was 3 units more negative.

Prevalent cancer model

For NSCLC and SCLC the probability of having cancer at baseline is given by a multinomial model with a histology-dependent slope term for PLCO (logit scale) and a stage-dependent intercept term. We have constrained the slope terms for PLCO to be non-negative (a greater smoking history cannot be associated with a lower risk of lung cancer at baseline). If the slope term is $\ln x$, then there is an odds ratio of x associated with a unit increase in logit-PLCO. Since the baseline risk is low, this means the relative risk is also approximately x for a unit increase. A unit increase in logit-PLCO represents, for example, the difference between smoking 10 cigarettes per day for 20 years and smoking 22 cigarettes per day for 20 years. If the slope term is 1 then the odds ratio in the prevalent cancer model will match the odds ratio in the original PLCO model for the probability of lung cancer in 6 years. We considered it likely that the slope term should be near 1 for this reason. We therefore used a log-normal distribution for each PLCO slope with parameters $\sigma^2=\ln 2$ and $\mu=-(\ln 2)/2$ in order to have mean and variance 1.

The further the slope parameter is from 0, the more the average risk of having cancer at baseline is dominated by those with highest PLCO, and we would ideally specify prior

distributions on the intercept parameters conditional on the slope parameters for PLCO, but since our aim was to include only weakly informative priors, we avoided this level of complexity. Instead, we adopted normal distributions for the intercepts with mean −8 and standard deviation 2.

Given this, the prior distributions for the PLCO slopes, and the data distribution of PLCO values, we obtain a prior predictive distribution (the distribution of an outcome of interest when parameters are drawn from the prior distribution) for the probability of baseline lung cancer which can approximately be characterised by a logit-normal distribution with parameters $\mu = -4$ and $\sigma = 1$, having expected value 0.032 and 80% of the distribution between 0.005 and 0.074. Since there are seven NSCLC stages in the model and two SCLC stages, this suggests that 7/9 of lung cancers are NSCLC and 2/9 are SCLC in the prior predictive distribution.

Preclinical NSCLC progression and clinical presentation

It is expected that some lung cancers will develop very quickly while others will take a considerable period of time to lead to symptoms and mortality. This is on the basis that late stage interval cancers are observed with short times between screening (e.g., in NLST and NELSON), evidence on volume doubling times (from NELSON) and that there is some evidence of overdiagnosis with LDCT screening. We aimed to use weakly informative priors for these parameters, and took inspiration from the mean preclinical sojourn time estimates from ten Haaf et al(25). After substantial algebraic manipulation to account for the different assumptions in the models (e.g., different stages used as model states), assuming zero heterogeneity in progression rates between cancers, and optimising using least squares, we produced estimates as shown in Table 5.

Table 5 Prior distributions for NSCLC progression and clinical presentation

Stage	λ		ξ	
	μ	σ	μ	σ
IA1, IA2, IA3	0.281	0.472	-2.608	0.472
IB	0.207	0.472	-1.786	0.472
II	0.592	0.472	-1.944	0.472
III	-0.268	0.472	-0.937	0.472
IV	N/A	N/A	0.190	0.472

With heterogeneity the calculations would become more difficult, but we assume they are a fair first approximation.

NSCLC heterogeneity

We looked to data on volume-doubling times in nodules which were left in situ and imaged multiple times in the NELSON study before being diagnosed as lung cancer(26). This may have had the effect of selecting out the most aggressive nodules (since it is unlikely these would be imaged when small enough to be left in situ). In this study the median volume-doubling time was 348 days, and the interquartile range was 222 to 492 days. Inverting these (so they are rates of doubling per year) we get a median of 1.05 and interquartile range 0.74 to 1.64. A log-normal distribution with similar quantiles (identified through least squares optimisation) is $\mathrm{LogN}(0.0777, 0.609)$, which has median 1.08 and interquartile range 0.72 to 1.63. We therefore consider that 0.609 is a plausible value for the heterogeneity in NSCLC progression rates. From this we construct a prior distribution for heterogeneity which is $\mathrm{LogN}(-0.5,1)$. This has median 0.607 and 80% of the distribution lies between 0.168 and 2.185.

SCLC progression and presentation

We followed a similar approach as for NSCLC and assigned a prior for the rate of progression from limited to extensive SCLC of $\mathrm{LogN}(1.45,0.472)$, and priors for the rates of presentation of $\mathrm{LogN}(2.07,0.472)$ for limited SCLC and $\mathrm{LogN}(0.562,0.472)$.

Context-specific presentation rates

We include additional terms for the rates of lung cancer presentation which are context-specific and added to the rates described above. The two contexts considered are mid-screening in NLST and post-screening in NLST. These rates are non-negative but we expect/prefer one or both of them to be near zero. We use prior distributions of $\mathrm{LogN}(-4,2)$.

Death from never diagnosed lung cancer

A prior distribution was applied to the proportion of patients who reach Stage IV NSCLC or extensive SCLC and die from the disease without being diagnosed (in the absence of other-cause mortality). The proportion is given by the transformation $f(\lambda) = \lambda/(\lambda + \xi)$ where λ is the rate of death from undiagnosed disease and ξ is the rate of clinical presentation. This has derivative $f'(\lambda) = \xi/(\lambda + \xi)^2$ and this is included as a Jacobian transform.

$$\frac{\lambda}{\lambda + \xi} \sim \mathrm{Beta}(1.0096,529.5)$$

was used, such that 99% of the prior distribution was between 10^{-5} and 10^{-2}, i.e., between 1 in 100,000 and 1 in 100 metastatic cancers would cause death but never be diagnosed.

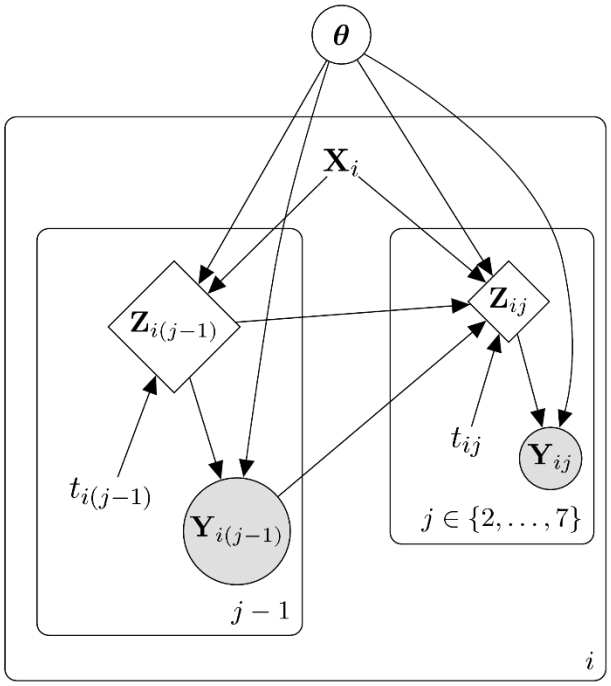
Likelihood function

The likelihood function is composed of a likelihood component for each NLST participant, i.e., the log-likelihood is the sum of the patient-level log-likelihoods. The patient-level likelihood function involves estimating and projecting a state vector which corresponds to the probability distribution of the true state of the patient at a given time over a set of possible health states. The possible health states include having no lung cancer, having preclinical NSCLC, having clinical NSCLC (i.e., diagnosed following symptomatic presentation or incidental diagnosis), having preclinical SCLC, having clinical SCLC, having died from undiagnosed NSCLC, having died from undiagnosed SCLC, and having died from other causes.

The observations of the patient at various points in time give incomplete information. We can observe that a patient is alive and has not been diagnosed with lung cancer (this rules out the clinical lung cancer states and the death states), we can observe that a patient has undergone a LDCT screen with a negative result (which pushes the state vector towards the no lung cancer state and early lung cancer states due to LDCT having higher sensitivity for more advanced lung cancer stages). We assume that observations of death or cancer diagnosis give perfect information. This approach has similarities to hidden Markov modelling (but in continuous time).

The initial distribution over the states is determined using a multinomial logistic model and is dependent on the PLCO linear predictor for the patient (see Section **Prevalence** above). Thereafter the model processes longitudinal data points for each patient, at each point potentially updating the likelihood and state. Each record is either a screening record or a follow-up record. A screening record refers to a specific point in time when a patient underwent a screen which is either negative or positive for lung cancer – if it is positive then the histology (NSCLC or SCLC) and stage are also noted. A follow-up record refers to a period of time ending in either the patient being: observed alive and without any lung cancer diagnosis; alive and diagnosed with lung cancer; dead without a prior lung cancer diagnosis (could be caused by death from other causes or by a never-diagnosed lung cancer).

Figure 2 Bayesian representation of the likelihood function calculation



Unshaded circles are latent parameters, shaded circles are statistically modelled observations, quantities without shapes are data which are not statistically modelled, diamonds are deterministically calculated quantities

In Figure 2 we show the factorisation of the likelihood. $\mathbf{Y}_{ij}$ are the observed outcomes for participant i in record j, and the log-likelihood is given as the joint log-probability of all the records for each participant. The joint log-probability is decomposed into a conditional probability through the use of state vectors $\mathbf{Z}_{ij}$ which are deterministically calculated and give the probability that an individual is in any particular health state at the time t_{ij} (when $\mathbf{Y}_{ij}$ is observed). The participant must either have no lung cancer, preclinical lung cancer (of a particular histology and stage), clinically identified lung cancer (of a particular histology and stage), or have died (from non-lung cancer causes or from a lung cancer that will never be diagnosed). For example, if the observed outcome in a record is a negative lung cancer screen, then $\mathbf{Z}_{ij}$ will be used (along with the test sensitivity estimate from $\boldsymbol{\theta}$) to calculate the pre-test probability of a negative screen.

The calculation of the log-likelihood proceeds in an iterative manner using an accumulator variable which is initially set to zero. Each record is processed in turn, and each record will typically update the log-likelihood accumulator and the state vector. Once all records for a participant are processed, the value of the accumulator is returned as the log-likelihood contribution from that participant.

Finally, the above procedure is in fact repeated a total of seven times per participant, with a different value of ω_i (the random effect for NSCLC progression heterogeneity) being used. These values are chosen and combined using a Gauss–Hermite quadrature rule in order to marginalise over the distribution of ω_i. Due to the computational complexity of the model, and the relatively high number of parameters (52) it was not possible to use the full likelihood calculation within the Stan sampler. Instead, we developed a surrogate for the log-likelihood function using a multivariate normal approximation and a Gaussian process to model the residual after the approximation.

Screening results

If the screening result is negative, we first calculate the probability of a negative screen result conditional on the state vector. The probability of a negative screen in the no lung cancer state is 1 (i.e., we assume 100% specificity within the model as we only include confirmed lung cancer diagnoses in the data). The probability of a negative screen in a preclinical lung cancer state is given by one minus the relevant sensitivity for the state, e.g., if the sensitivity for Stage III NSCLC is 95%, then the probability of a negative screen conditional on having Stage III preclinical NSCLC is 0.05. The overall probability of a negative screen is the weighted average across the state vector of the conditional probabilities.

The probability of a negative screen is contributed to the likelihood function (i.e., the negative screen is the data, and the likelihood function includes the probability of observing that data given the parameters). The state vector is finally updated so that for each state we now have the conditional probability the patient is in the state given that they had a negative screen, e.g., for the no lung cancer state, we have a conditional probability as given in Equation (2).

$$\begin{aligned}&\Pr(\text{No lung cancer} \mid \text{Negative screen})\\&\quad= \frac{\Pr(\text{Negative screen} \mid \text{No lung cancer})\Pr(\text{No lung cancer})}{\Pr(\text{Negative screen})}\\&\quad= \frac{\Pr(\text{No lung cancer})}{\Pr(\text{Negative screen})}\end{aligned} \qquad (2)$$

If the screening result is positive, the likelihood function is instead updated with the appropriate contribution, which is the product of the probability of being in the relevant state (e.g., if the screening result is positive Stage II NSCLC then we refer to the probability of having Stage II preclinical NSCLC) and the relevant sensitivity. There are no updates to the state vector because there will be no subsequent records to process.

Follow-up

For a follow-up record, we first project the state vector forwards in time according to the various risks in the natural history model. These are represented by a matrix differential equation as shown in Equation (3) , where $\mathbf{y}(t)$ is the state vector at time t and $A(t)$ is a matrix containing the instantaneous hazards of transitions between different states.

$$\mathbf{y}'(t) = A(t)\mathbf{y}(t) \quad (3)$$

After projecting the state forward to the time at the end of the record, we then process the record depending on the event recorded. If the patient survives to the end of the period without dying or being diagnosed with lung cancer, we update the likelihood function with the probability of being in the no lung cancer or the preclinical lung cancer states. We then update the state vector accordingly (since the probability of being in the clinical lung cancer or death states is 0 at the end of the time period).

If the patient is diagnosed with lung cancer at the end of the period or dies at the end of the period, we use the derivative with respect to time of the relevant state as the likelihood contribution. E.g., if $y_i(t)$, $y_j(t)$ and $y_k(t)$ are the probabilities of being in the dead from undiagnosed NSCLC, dead from undiagnosed SCLC, and dead from other causes, then for a record ending in death without any lung cancer diagnosis we record the likelihood contribution as $y_i'(t) + y_j'(t) + y_k'(t)$. We do not update the state vector because no further records will exist for the patient (they have left the scope of the model).

Calibration and validation procedure

Surrogate for the log-likelihood function

The method described above for calculating the likelihood contribution for each patient is extremely computationally intensive. It generally requires dozens of ordinary differential equations to be solved for each patient each time the likelihood is calculated (and automated differentiation of this to also calculate the likelihood gradient). In order to facilitate model fitting within a practical timeframe we developed a surrogate (or emulator) for the likelihood function which was much more computationally tractable.

We chose to use Gaussian process regression since it is capable of incorporating high orders of interaction between parameters (which is more challenging with alternative methods such as generalised additive models). Furthermore, we used a multivariate normal distribution as a first approximation of the log-likelihood function, and used the Gaussian

process regression to fit the residual (the difference between the multivariate normal approximation and the true log-likelihood function). For both the multivariate normal approximation and the Gaussian process, it was important to have good quality training points.
The natural history model has 52 parameters (including 16 relating to LDCT and CXR screening), with the result that the typical set may occupy an incredibly small volume of the "total" parameter space (in the sense that a randomly selected combination of parameter values is almost certain to lie outside the target set).

Another result of the high dimensionality of the model is that the mode (if there is exactly one) of the posterior density (the maximum a posteriori estimate) is not necessarily in the typical set. As is described in Figure 3 of *A Conceptual Introduction to Hamiltonian Monte Carlo*, the typical set is expected to be some distance away from the mode(27). Indeed, "plausible parameter values" (i.e., in the typical set) may have a surprisingly low posterior density function in comparison with the maximum a posteriori estimate.
We first generated a set of 10,000 parameter sets using Latin hypercube sampling and hand-tuned transformations from the unit hypercube to the support of the parameters. We also included the maximum a posteriori estimate for parameters which was obtained through a standard optimisation approach. We calculated the log-posterior density for each of these parameter sets and trained a random forest classifier to classify points as above or below a log-posterior threshold, and then used a slice sampler to sample 10,000 new points which were expected to have a log-posterior higher than the threshold according to the random forest classifier. This process was repeated, with the threshold for each iteration set by hand, until no further improvements were achieved (after five iterations).

Next, we attempted to fit a multivariate normal approximation to the log-posterior values obtained in previous iterations by finding the least-squares solution to

$$y_i = c - (\mathbf{x}_i - \mathbf{x}^\star)^\top \mathrm{A} (\mathbf{x}_i - \mathbf{x}^\star) + \epsilon_i$$

where A is a positive definite matrix, $\mathbf{x}^\star$ is a central point, and c is an intercept term. We excluded points where the log-posterior was extremely low because these were very unlikely to be in the typical set. We then used this to sample further parameter sets and evaluate the log-posterior.

After several further iterations, we fitted a multivariate normal approximation to the log-likelihood, and additionally fitted a Gaussian progress to the residual error. The initial values for c, A and $\mathbf{x}^\star$ were based on the maximum likelihood estimate (MLE) of parameters, the log-likelihood evaluated at the MLE estimate, and the Hessian matrix (the matrix of second partial derivatives) of the log-likelihood estimated using a finite differences approach. The

Gaussian process used a radial basis function kernel with automatic relevance detection (i.e., separate lengthscales for each parameter).

This combined model was used within Stan to sample from the posterior distribution. The model was run with four chains, each for 2,000 iterations with the first 1,000 iterations discarded. Initial values were parameter sets with high log-likelihood from previous rounds. For each set of posterior samples we then evaluated the true log-likelihood, refitted the surrogate, and repeated. We stopped when we judged there was good concordance between the surrogate and the true log-likelihood. The final surrogate used a Gaussian process with 12,000 training points.

Calibration

We ran the Stan sampler using the surrogate for the log-likelihood function. The model was run with four chains, each for 2,000 iterations with the first 1,000 iterations discarded. Initial values were parameter sets with high log-likelihood. We verified that there were no divergent transitions and verified convergence by assessing the $\hat{R}$ diagnostic and confirming it was equal to 1. We examined energy plots(27), and examined effective sample sizes (considering an effective sample size to actual sample size ratio of <0.1 indicative of an issue).

Validation

We validated against NLST by constructing a set of aggregate results from each sample of the posterior distribution. For each sample of the posterior distribution, for each participant, we used their $PLCO_{m2012}$ value, age and current smoking status to simulate the probability over time that they are in particular health states and that particular observations would be made in the following categories:

- Cancers detected before a baseline screen
- Cancers detected at baseline (T0) screen (number and stage)
- Interval cancers diagnosed between the baseline screen and T1 screen
- Cancers detected at the T1 screen
- Interval cancers diagnosed between the T1 and T2 screen
- Cancers detected at the T2 screen
- Cancers diagnosed within three years of the T2 screen.

The mean, 5^{th} and 95^{th} percentiles of the posterior distributions of these predictions were compared to the actual numbers from NLST as well as to a non-parametric bootstrap of NLST (to indicate how much uncertainty is associated with each number in NLST). These simulations needed to account for certain factors which do not need to be included in the

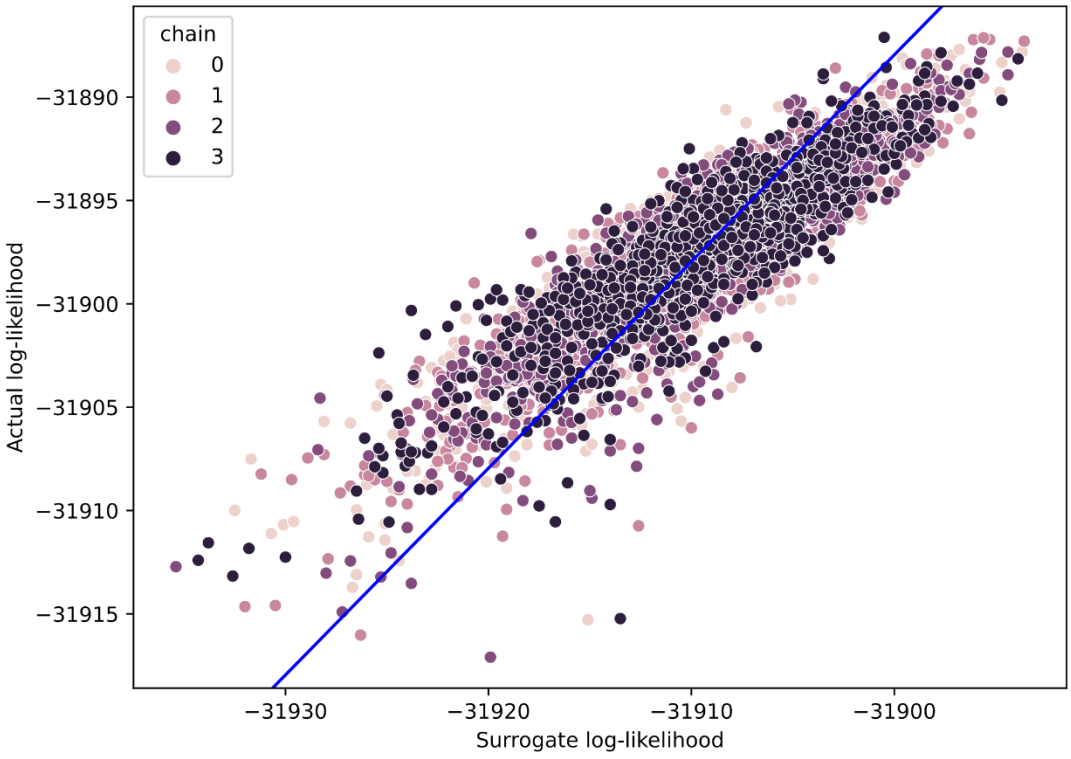
natural history model itself, e.g., missed screens (and the association between smoking history and missing a screen). We modelled many of these factors, but we did not investigate whether any of them varied according to the trial arm (i.e., performance bias). There may be other factors that are not modelled, e.g., informative censoring (participants may be more likely to be considered lost to follow-up if they never get diagnosed with lung cancer).

Results

Surrogate log-likelihood

The surrogate log-likelihood function was a fairly good proxy for the actual log-likelihood function, as shown in Figure 3. The blue line in the figure has slope 1 and goes through the mean of the two functions; for a perfect surrogate all points would be on the blue line. It is not important that the surrogate is slightly shifted compared to the actual log-likelihood – additive constants are irrelevant for sampling in Stan. The surrogate is slightly underestimating the range of log-likelihood and there is clearly some error/noise.

Figure 3 Comparison of surrogate log-likelihood and actual log-likelihood



Future work may involve using a statistical technique (e.g., independence chain Metropolis–Hastings) to adjust for the difference between the surrogate log-likelihood and the actual log-likelihood to remove any bias introduced by using the surrogate.

Posterior distribution

A summary of the posterior distribution is given in Table 6. Samples from the posterior distribution have been made open access.

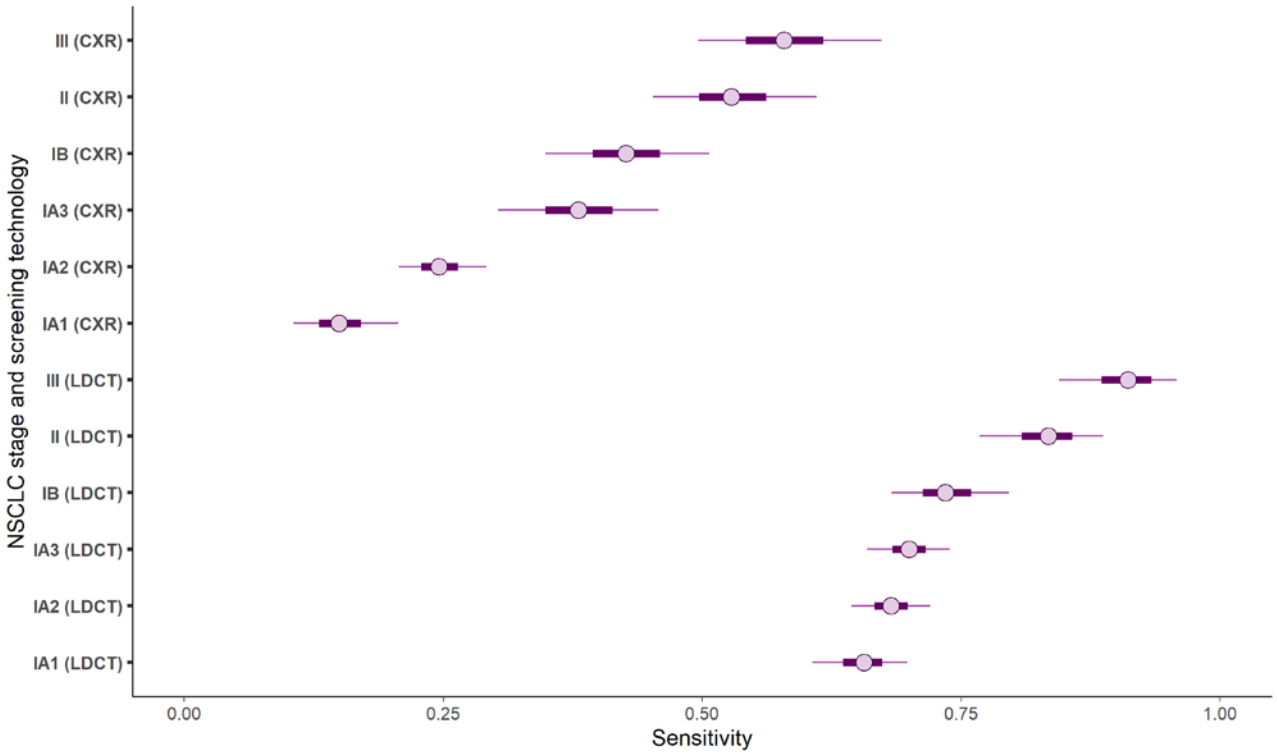
Table 6 Posterior distribution summary

Parameter	Mean (90% credible interval)
Sensitivity of testing technology for NSCLC (log-odds scale)	
Low-dose CT	
Stage IA1	0.641 (0.432, 0.838)
Stage IA2	0.766 (0.593, 0.944)
Stage IA3	0.848 (0.660, 1.040)
Stage IB	1.037 (0.766, 1.363)
Stage II	1.619 (1.195, 2.060)
Stage III	2.358 (1.691, 3.122)
Chest X-ray	
Stage IA1	-1.742 (-2.138, -1.345)
Stage IA2	-1.119 (-1.344, -0.887)
Stage IA3	-0.490 (-0.833, -0.171)
Stage IB	-0.297 (-0.624, 0.027)
Stage II	0.121 (-0.190, 0.449)
Stage III	0.331 (-0.017, 0.722)
Sensitivity of testing technology for limited SCLC	
Low-dose CT	0.598 (0.403, 0.791)
Chest X-ray	0.227 (0.136, 0.338)
Final screen effect (log-odds scale)	
Low-dose CT	0.898 (0.333, 1.808)
Chest X-ray	0.058 (0.021, 0.119)
NSCLC prevalence	
PLCO coefficient	0.887 (0.811, 0.965)
Stage-specific intercept	
Stage IA1	-7.128 (-7.426, -6.835)
Stage IA2	-5.663 (-5.794, -5.533)
Stage IA3	-6.925 (-7.218, -6.643)
Stage IB	-11.178 (-12.326, -10.068)
Stage II	-6.839 (-7.048, -6.631)

Parameter	Mean (90% credible interval)
Stage III	-6.563 (-6.719, -6.401)
Stage IV	-7.163 (-7.359, -6.965)
SCLC prevalence	
PLCO coefficient	0.939 (0.739, 1.168)
Stage-specific intercept	
Limited	-7.794 (-8.251, -7.338)
Extensive	-8.259 (-8.587, -7.930)
NSCLC incidence	
Parameter ‘c’	-5.572 (-5.632, -5.513)
Parameter ‘m’	0.839 (0.787, 0.892)
SCLC incidence	
Parameter ‘c’	-7.486 (-7.599, -7.372)
Parameter ‘m’	0.979 (0.876, 1.083)
NSCLC progression rates (per year)	
Stage IA1	6.389 (5.362, 7.507)
Stage IA2	1.555 (1.336, 1.805)
Stage IA3	5.015 (4.217, 5.886)
Stage IB	11.428 (9.308, 13.748)
Stage II	3.124 (2.638, 3.656)
Stage III	1.100 (0.946, 1.272)
Death from undiagnosed Stage IV	0.004 (0.001, 0.010)
Random effects parameter σ	0.935 (0.826, 1.051)
NSCLC presentation rates (per year)	
Stage IA1	0.022 (0.014, 0.033)
Stage IA2	0.040 (0.029, 0.054)
Stage IA3	0.083 (0.060, 0.111)
Stage IB	0.267 (0.216, 0.323)
Stage II	0.288 (0.239, 0.344)
Stage III	0.472 (0.402, 0.551)
Stage IV	1.939 (1.648, 2.266)
SCLC progression rates (per year)	
Limited	0.621 (0.470, 0.792)
Death from undiagnosed extensive	0.004 (0.001, 0.010)
SCLC presentation rates (per year)	
Limited	0.199 (0.133, 0.285)
Extensive	2.128 (1.579, 2.791)
Context-specific presentation rates	
Mid-screening	0.002 (0.001, 0.005)
Post-screening	0.064 (0.050, 0.081)

The sensitivity of LDCT and CXR (without the final screen effect) for NSCLC are shown in Figure 4. We predict that LDCT is substantially more sensitive than CXR, and this is identified purely from the data, since the priors for sensitivity were the same for LDCT and CXR.

Figure 4 Sensitivity of low-dose CT and chest X-ray for NSCLC in the natural history model



We examined the effect of PLCO on prevalence and incidence by evaluating these for the 5th, 25th, 50th, 75th and 95th centile of PLCO in NLST, as shown in Table 7. There is a clear pattern that prevalence and incidence both increase in line with the PLCO predicted risk. We also predict that incidence is lower for ex-smokers than for current smokers, even if they have the same current PLCO score (note that the PLCO score is lowered by quitting smoking).

Table 7 Prevalence and incidence of lung cancer based on the natural history model

Outcome	$PLCO_{m2012}$ [predicted 6-year risk, centile in NLST]				
	0.8%, 5th	1.5%, 25th	2.4%, 50th	4.0%, 75th	8.7%, 95th
Probability of prevalent disease [per 100,000 (mean, 90% credible interval)]					
Any disease	331 (287, 376)	572 (516, 629)	883 (813, 956)	1420 (1320, 1530)	2920 (2630, 3210)
NSCLC					
Stage IA1	29.1 (20.6, 39.1)	50.3 (36.4, 66.9)	77.6 (56.7, 102)	125 (91.5, 164)	255 (186, 338)
Stage IA2	124 (105, 147)	215 (186, 247)	332 (290, 376)	532 (466, 604)	1090 (931, 1260)
Stage IA3	35.7 (25.5, 47.7)	61.6 (45.0, 81.2)	95.0 (69.7, 124)	152 (113, 198)	312 (230, 407)
Stage IB	0.6 (0.2, 1.5)	1.1 (0.3, 2.6)	1.7 (0.4, 4.0)	2.7 (0.7, 6.5)	5.6 (1.4, 13.4)
Stage II	38.5 (30.4, 48.5)	66.6 (53.3, 82.1)	103 (82.5, 125)	165 (133, 201)	337 (270, 415)
Stage III	50.7 (40.9, 61.6)	87.6 (73.0, 104)	135 (115, 158)	217 (186, 251)	443 (376, 515)
Stage IV	27.8 (22.0, 34.6)	48.1 (38.9, 58.6)	74.2 (60.5, 90.0)	119 (97.7, 144)	244 (196, 299)
SCLC					
Limited	14.6 (7.93, 23.6)	25.8 (15.3, 40.0)	40.6 (24.7, 61.8)	67.1 (41.3, 101)	144 (84.5, 225)
Extensive	9.0 (5.4, 13.7)	15.9 (10.6, 22.5)	25.1 (17.6, 34.3)	41.3 (29.7, 55.4)	88.7 (60.7, 125)
Probability of incident preclinical disease within 5 years [per 100,000 (mean, 90% credible interval)]					
NSCLC					
Current smoker	909 (828, 992)	1520 (1430, 1630)	2300 (2170, 2420)	3590 (3400, 3780)	7040 (6520, 7560)
Ex-smoker	792 (716, 871)	1330 (1235, 1426)	2000 (1890, 2120)	3130 (2970, 3300)	6150 (5730, 6570)
SCLC					
Current smoker	120 (99.9, 143)	219 (193, 249)	355 (320, 392)	601 (546, 658)	1340 (1160, 1530)
Ex-smoker	102 (83.5, 123)	187 (162, 215)	301 (270, 336)	511 (465, 558)	1140 (999, 1290)

All numbers are given to a maximum of three significant figures and one decimal place; probabilities of incident preclinical disease ignore competing risks

We examined the posterior distribution for the stage distribution of NSCLC in the absence of screening, ignoring the competing risk of death. As shown in Table 8, the stage distribution is heavily dependent on how aggressive a cancer is. The least aggressive cancers stand a fair chance of being detected while in a fairly early stage, while the most aggressive cancers are extremely likely to present in Stage III or IV.

Table 8 Stage distribution of NSCLC in the absence of screening

Stage	Centile of random-effects distribution for heterogeneity in progression rates				
[% (mean, 90% credible interval)]	5th	25th	50th	75th	95th
IA1	2.5 (1.5, 3.6)	1.0 (0.7, 1.4)	0.5 (0.4, 0.8)	0.3 (0.2, 0.4)	0.1 (0.1, 0.2)
IA2	15.2 (11.1, 20.2)	6.9 (5.1, 9.1)	3.8 (2.8, 5.0)	2.1 (1.5, 2.7)	0.8 (0.6, 1.2)
IA3	8.9 (6.3, 12.0)	4.3 (3.1, 5.7)	2.4 (1.8, 3.2)	1.3 (1.0, 1.7)	0.5 (0.4, 0.7)
IB	10.6 (8.3, 13.4)	5.6 (4.4, 7.0)	3.3 (2.6, 4.0)	1.8 (1.5, 2.2)	0.8 (0.6, 1.0)
II	25.0 (22.2, 27.8)	17.5 (14.9, 20.2)	11.3 (9.7, 13.0)	6.7 (5.8, 7.7)	2.9 (2.4, 3.4)
III	28.4 (23.0, 33.1)	35.9 (33.2, 38.5)	31.4 (28.6, 34.6)	23.0 (20.8, 25.3)	11.9 (10.6, 13.3)
IV	9.5 (5.4, 14.1)	28.9 (23.0, 34.5)	47.2 (42.5, 51.5)	64.8 (61.9, 67.5)	82.9 (81.0, 84.7)

Checking and validation

Diagnostic tests suggested no issues with the model. There were no divergent transitions. $\hat{R}$ was 1 for all parameters (indicating convergence between chains). Energy plots suggested no difficulty with heavy tails in the posterior distribution. The ratio of effective sample size to actual sample size was between 0.5 and 1 for all but one parameter (the rate of progression from preclinical limited SCLC to preclinical extensive SCLC), and for this it was 0.48.

There was generally good agreement between the posterior predictive distributions and the bootstrap distributions for NLST, as shown in Figure 21 (Appendix 1). Nevertheless there were some occasions where the natural history model predicts somewhat different outcomes, for example:

- The number of cancers detected at the second (T1) screen in the CXR arm are higher in the posterior predictive distribution from the model than in NLST, while the number of cancers detected in the interval between the T1 and T2 screens in the CXR arm are lower than in NLST – this could be consistent with the sensitivity of CXR being overestimated in the model, or an underestimation of how many participants miss the T1 screen in the CXR arm;
- The number of late-stage (Stage III/IV) cancers in the final (T2) screen in the LDCT arm are lower in the posterior predictive distribution from the natural history model than in NLST, while the number of late-stage cancers in the T1–T2 interval and post-screening is higher;
- The number of Stage IA2 cancers in both arms are lower in the posterior predictive distribution from the natural history model than in NLST, which could (for example) reflect some participants receiving an additional screen beyond the trial protocol.

Overall, the natural history model replicates many of the findings of NLST, e.g., significantly more cancers diagnosed in LDCT screening rounds than CXR screening, significantly more interval cancers in the CXR arm, relatively high abundance of Stage IA2 NSCLC at screening rounds, relatively low abundance of Stage IB NSCLC at screening rounds, relatively high numbers of late-stage cancers diagnosed in the intervals between screens and after screening finished.

Economic evaluation of targeted lung cancer screening

Modelling approach

Perspective, time horizon and discounting

The model perspective was that of the NHS and Personal Social Services. The direct effects of individuals contacted through the screening programme were included.
A life-time horizon was taken, with most simulated individuals having died before age 100.
Costs and health outcomes both discounted at 3.5% per annum(28).

Analysis method

A cost-utility analysis was undertaken, where the costs and quality-adjusted life years (QALYs) were estimated and compared for each of the 49 strategies. Based on the incremental cost-effectiveness ratio (ICER, the incremental costs divided by incremental QALYs), strategies that are dominated (i.e. their incremental QALYs are lower and incremental costs higher than one or more other strategies), or extendedly dominated (i.e. their ICER is greater than that of the next more effective strategy) are eliminated, and a cost-effectiveness frontier is created. In the Results section, only findings relevant to strategies on the cost-effectiveness frontier are reported.

Although a fully incremental cost-effectiveness analysis (where dominated and extendedly dominated options are eliminated) is appropriate, there are numerous deficiencies with ICERs, so we additionally calculate the *net benefit*(29). The net benefit of an option is the health benefits obtained by pursuing the option, minus the cost incurred by pursuing the option. Although costs and benefits are not on the same scale, they can be placed on a common scale according to willingness-to-pay (WTP) or the opportunity cost of health spending. The net monetary benefit (NMB) is obtained when health benefits are multiplied by WTP before costs are subtracted, while net health benefit (NHB) is obtained when costs are divided by WTP before being subtracted. The option which gives the greatest net benefit (NMB or NHB) is the economically optimal choice.

Results from all analyses were reported in terms of the incremental net monetary benefit (INMB). The INMB is calculated using incremental costs and QALYs for one strategy compared with another (in this report, always the no screening strategy), assuming a specific WTP threshold. The WTP threshold reflects the monetary value of the QALYs gained. For instance, the National Institute for Health and Care Excellence (NICE) state that their WTP threshold is between £20,000 per QALY gained and £30,000 per QALY gained.

The IMNB is calculated as

(Incremental QALYs * WTP threshold) – Incremental costs.

A positive INMB suggests that the strategy is cost-effective compared to the alternative strategy at the specific WTP threshold. A negative INMB would indicate that the strategy is not cost-effective compared to the alternative at this WTP threshold. In the following results we applied the WTP thresholds as quoted by NICE: £20,000 per QALY gained and £30,000 per QALY gained. Using the INMB helps to more easily determine which is the optimal screening strategy, i.e. the strategy providing the greatest INMB at the different WTP thresholds.

The main analysis was a probabilistic analysis, where parameter values were sampled from relevant distributions to reflect parameter uncertainty. A microsimulation analysis, where all parameters were set at their central values, is also presented. Controlled one-way sensitivity analyses were undertaken to assess the impact of increasing and decreasing the value of each parameter in turn by 20%. Additional scenario analyses were conducted to evaluate the impact of specific parameters or assumptions.

Software

The DES model was developed in Microsoft Excel (Microsoft Corporation). Additional analyses for updated parameters were conducted in Stata 16.0 (StataCorp LP, College Station, TX, USA).

Overall model structure

No changes were made to the general modelling approach as described in Snowsill 2018(5). A cohort of individuals was simulated with a range of baseline characteristics (including age and predicted risk of lung cancer). Each individual was concurrently simulated with four screening intervention arms and the no screening arm. By simulating the same individuals concurrently through all arms there is a reduction in stochastic variation. The costs, QALYs and other outcomes for each full programme (combination of targeting strategy and intervention) were estimated using a decision tree. Costs of

administering the screening programme were accumulated through the decision tree, and long-term costs and QALYs were estimated at the leaves of the decision tree by identifying appropriate individuals simulated in the cohort and assigning them appropriately either to the screening intervention (if they meet all criteria and join the screening programme) or to no screening(5).

The DES modelling involved sampling times to future events according to the current state of the individual (and any relevant history). The earliest of these events was modelled as occurring and the model 'clock' advances to that event. Times to events were then either reduced by the amount the clock has advanced or were resampled (as appropriate) (5). Certain event times were simulated identically for each individual across the strategies under consideration, e.g., the time from the start of the model to when they would die from other causes (if they do not die from lung cancer beforehand). This further served to reduce stochastic variation.

In a change to the original model structure (see Figure 22 in Appendix 2), NSCLC and SCLC were explicitly modelled in this update with seven NSCLC stages (IA1, IA2, IA3, IB, II, III and IV, see Figure 5), and two SCLC stages (limited and extensive, see Figure 6).

To accommodate the revision of cost parameter estimates, a change was made to the structure of the original DES model (5). In order that newer innovative higher cost drugs were accounted at the right time, recurrent disease was explicitly modelled. The risk of the new event of 'Recurrence' was added to the decision logic for individuals with clinical lung cancer stage I to III. Recurrence was not allowed if stage IV was already reached. Allied to this, risk of death from lung cancer in stages I to III was annulled to create the requirement of passage through stage IV lung cancer prior to lung cancer death – unless other cause mortality occurs first. Instead, time to recurrence after diagnosis in stages I-III is estimated based on the sampled time to lung cancer death for stages I-III minus the sampled time to lung cancer death from stage IV. For example, time to recurrence from stage IIB is equal to the sampled time to death from stage IIB minus the sampled time to death from stage IV. See Figure 5 and Figure 6. As described in previous chapters, the outputs from the natural history model (including LDCT sensitivity, lung cancer prevalence, incidence and progression) were used as inputs for the DES model.

Figure 5 Model diagram for simulating individuals with NSCLC.

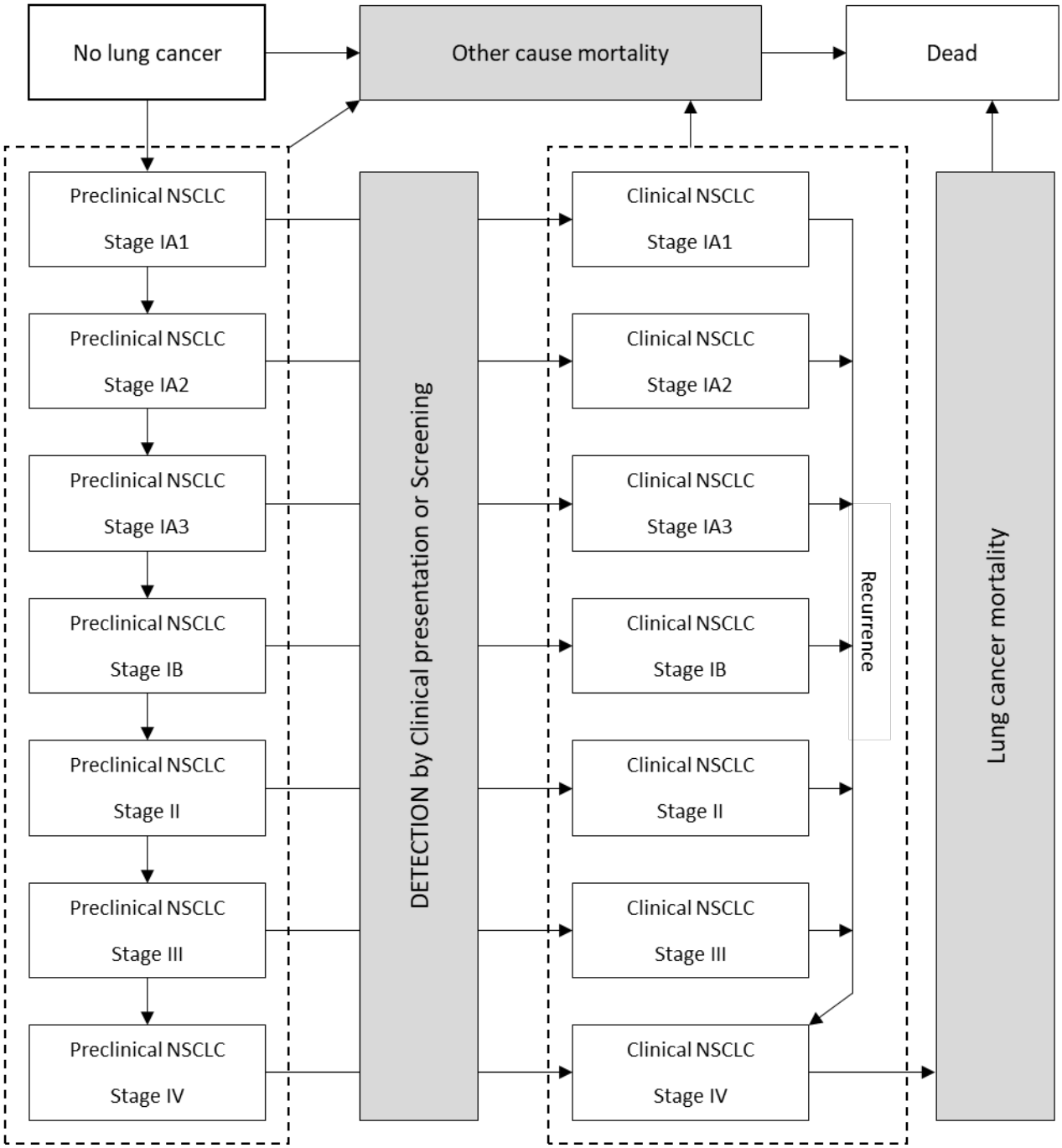
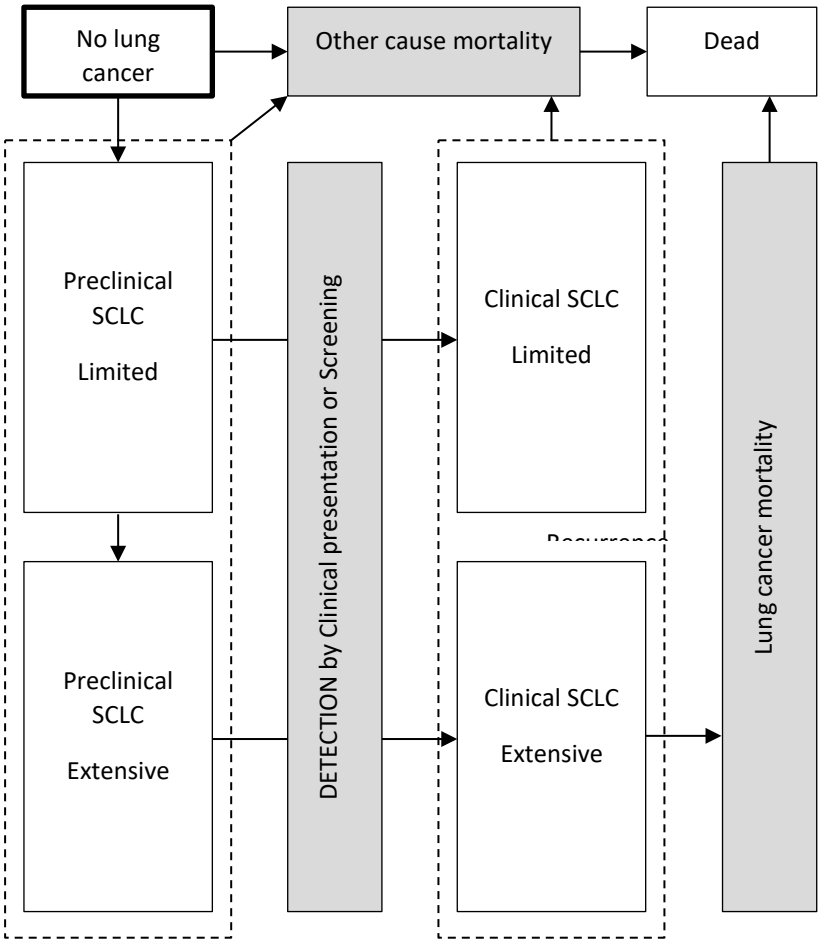


Figure 6 Model diagram for simulating individuals with SCLC



Model parameters

Mortality

Mortality after a lung cancer diagnosis

In our interim report to the NSC on the cost-effectiveness of LDCT to screen for lung cancer in high risk individuals, estimated survival after a lung cancer diagnosis was based on data from Goldstraw 2016(30) and adjusted to match UK estimates of one-year survival for individuals with PS 0-2. In that approach (similar to that used in the original report(5)) stage-specific survival estimates were applied to all lung cancers, regardless of whether they were screen-detected, interval cancers or cancers identified after or outside of the screening programme (i.e. in the comparator arm). This approach assumed that:

- there was no survival benefit if a screen-detected cancer was diagnosed in the same stage as it would have presented clinically (i.e. if a stage shift was not observed)
- there was no survival benefit in some cases when a stage shift was achieved (because the sampled survival is less than the counterfactual survival plus lead time)
- all cancers – regardless of how identified – were assumed to be of equivalent aggressiveness
- everyone who enters the screening programme had PS 0-2 at entry and throughout the model.

A further limitation of this approach, was the lack of survival data beyond 6 years from Goldstraw 2016(30). To account for these limitations, in the base case analysis, data from the NLST were used to estimate survival. Use of NLST allows explicit modelling of survival for screen-detected cancers as distinct from interval cancers and cancers diagnosed after/outside of the screening programme. This means that a modelled individual could see a benefit in survival from being screen-detected rather than presenting clinically even if they were detected in the same stage in which they would have presented. Survival for interval cancers has been found to be lower than survival for screen-detected cancers(31). This approach required no assumptions on PS in the estimation of survival.

Survival data from the screening arm of NLST were analysed separately for NSCLC and SCLC using the Stata streg command to fit generalised gamma models. The covariates were age, stage at diagnosis, and whether a cancer was detected by screening, and if it was not detected by screening, whether it was diagnosed at least 1000 days post-randomisation. Cancer diagnoses made after 1000 days post-randomisation were assumed to reflect cancers detected either post-screening or outside of the screening programme (including those in the no screening arm in the economic model). Resulting coefficient estimates from these analyses for NSCLC and SCLC are shown in Table 9 below.

Table 9 Coefficient estimates for survival after a lung cancer diagnosis from NLST

Coefficients		NSCLC estimate (variance)	SCLC estimate (variance)
Constant		1.276 (0.011)	1.151 (0.033)
Route of cancer detection	Screening	Reference group	Reference group
	Interval	-0.683 (0.017)	-0.142 (0.066)
	Post-/outside screening	-0.533 (0.010)	-0.145 (0.036)
Stage	IA1	2.215 (0.067)	NA
	IA2	1.921 (0.025)	NA
	IA3	1.821 (0.052)	NA
	IB	1.437 (0.042)	NA
	II	1.187 (0.026)	NA
	III	Reference group	NA

	IV	-1.382 (0.012)	NA
	Limited	NA	Reference group
	Extensive	NA	-1.413 (0.033)
Age*		-0.024 (0.001)	-0.045 (0.001)
ln(sigma)		0.311 (0.001)	0.186 (0.006)
kappa		0.307 (0.007)	0.604 (0.038)

* age is centred on 64 years

NA, not appropriate; NSCLC, non-small cell lung cancer; SCLC, small cell lung cancer

Figure 7, Figure 8 and Figure 23 (in Appendix 2) illustrate the estimated survival over time by stage at diagnosis and route of detection for individuals with NSCLC. Regardless of the stage, screen-detected lung cancers were estimated to have better survival than cancers detected outside of the screening programme. Interval cancers were associated with the lowest estimates of survival.

Figure 24 (in Appendix 2) illustrates estimated survival by SCLC stage at diagnosis and route of detection is presented. There is little differentiation in survival across the route of detection for SCLC.

Figure 7 Survival for a 70-year old individual diagnosed with stage IA1 NSCLC by route of detection

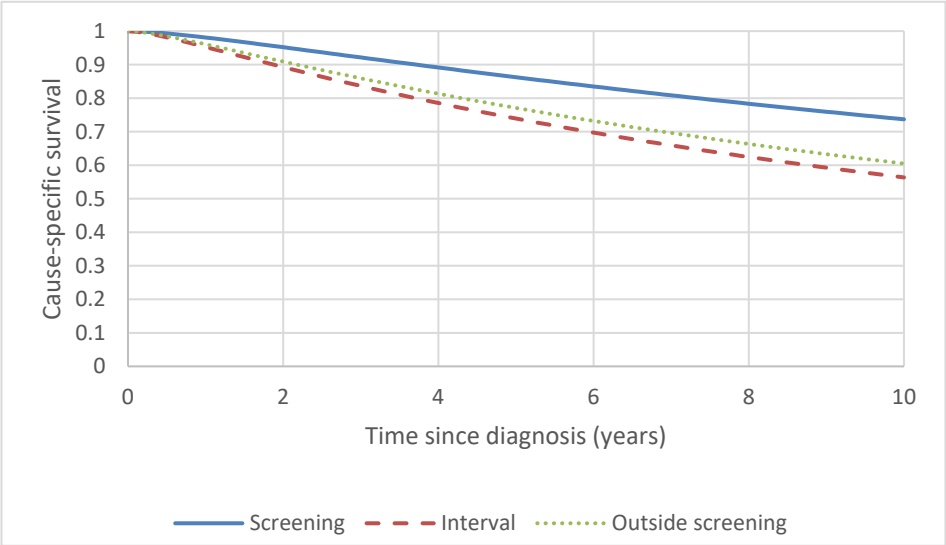
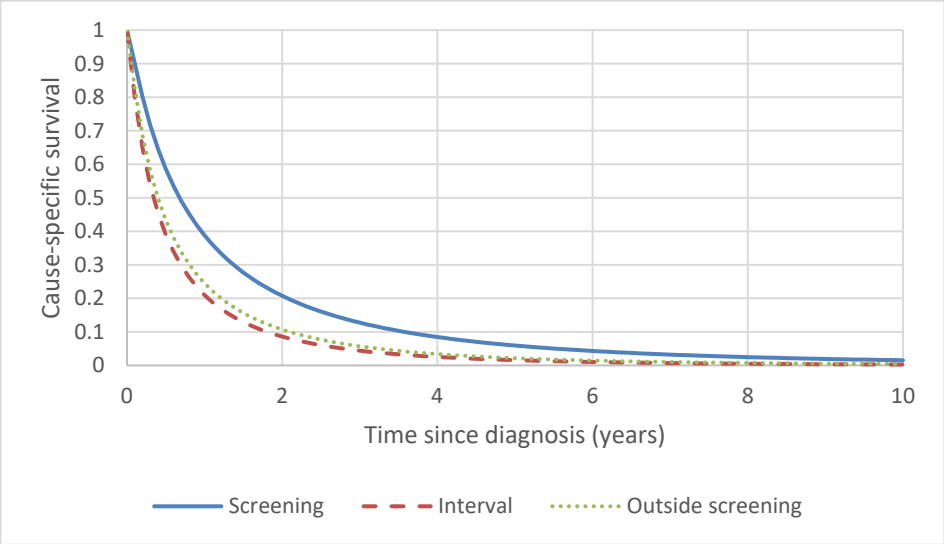


Figure 8 Survival for a 70-year old individual diagnosed with stage IV NSCLC by route of detection



To review the validity of this approach, estimated survival based on data from NLST were compared with limited data available from the UKLS (kindly supplied by David Baldwin and John Field, analysed by M Davies) and NLCA (kindly supplied by David Baldwin). Notwithstanding the fact that the different data sources use different stages and classification systems, this exercise provides a general picture on the face validity of using data from NLST.

Five-year survival estimates from NLST for stage IA were more favourable than those from UKLS (Figure 25 in Appendix 2). When diagnosis at stage IB was also considered, the survival estimates appeared more consistent (Figure 26 in Appendix 2).

When compared to NLCA data for PS 0-2, 1-year survival estimates from NLST for those cancers detected after/outside of screening were reasonably consistent (see Figure 27, Figure 28 and Figure 30 in Appendix 2). A slightly higher estimate of survival for stage II is seen from NLST compared to NLCA (Figure 29 in Appendix 2), and a much lower estimate is seen from NLST for stage IV (Figure 31 in Appendix 2). When compared to NLCA data for all PS groups, the NLST survival estimates are higher across all stages (not shown).

Thus, compared to the limited UK data available (5 year survival for diagnosis at stage I-II, based on 42 patients), use of NLST to estimate screen-detected lung cancers, may lead to the benefits of screening on survival being overestimated. For clinically presenting lung cancers, use of NLST data may underestimate survival in stage IV, but appears somewhat consistent with UK estimates for earlier stages. Considering these findings, use of the

NLST data to estimate lung cancer survival may overestimate the benefits of LDCT screening. To address this potential, two scenario analyses for calculating non-screen detected lung cancer survival were conducted.

In these scenario analyses data from Goldstraw 2016(30) are adjusted to fit 1-year survival estimates from NLCA data for non-screen-detected cancers. Thus, a limitation in these analyses was that interval cancers are not considered differently to cancers detected after/outside of the screening programme. The first sensitivity analysis assumed 1-year survival for individuals with PS 0-2. The second assumed 1-year survival for all individuals (i.e. across all PS groups, 0-4).

NSCLC survival data by stage (IA1, IA2, IA3, IB, IIA, IIB, IIIA, IIIB, IIIC, IVA, IVB) were extracted from Figure 2B of Goldstraw(30). Data were combined across sub-stages that were not explicitly modelled (e.g. IIA and IIB; IIIA, IIIB and IIIC; IVA and IVB). Survival was also estimated for individuals with SCLC by stage (limited, extensive) using Khakwani 2014(32), to align with the modelling of individuals with SCLC in the natural history model.

To estimate NSCLC survival, a proportional hazards Weibull model was found to be appropriate. In Stata, a weighted linear regression was performed on the log cumulative hazard, with log time and stage as independent variables. As in Snowsill 2018(5), weights were defined as the number of patients diagnosed within each stage multiplied by survival to approximate the number of patients contributing. Each survival curve was then described by a lambda parameter (specific for each stage) and a shared gamma parameter. To adjust these survival curves to reflect the lower survival estimates observed in the UK, the gamma parameter estimated from the weighted regression was assumed for all stages, and the lambda parameters estimated for all stages were adjusted to fit the 1-year survival estimates obtained from the NLCA. NLCA data for individuals with PS 0-2 were used in sensitivity analysis1, and NLCA data for individuals with PS 0-4 were used in sensitivity analysis 2.

A similar approach was taken for the estimate of SCLC survival. Kaplan-Meier survival data reported by Khakwani 2014(32) for individuals receiving radiotherapy and chemotherapy with limited and extensive SCLC from the NLCA were extracted. Again, a proportional hazards Weibull model was found to be appropriate, and a weighted linear regression was conducted. As the 1-year survival for extensive SCLC from the Khakwani 2014 data (from 2004 to 2011) reflected that from the updated NLCA 2017-2018 (39%, reflecting little improvement in survival over time for patients with extensive SCLC), no adjustment was made. However, given the more recent NLCA data indicated improved survival for patients with limited SCLC (75% vs 69%), the Weibull curve was adjusted to reflect this.

Mortality from undiagnosed lung cancer

It was assumed that there is no hazard of dying from lung cancer with undiagnosed lung cancer. However, once a patient is diagnosed their hazard of dying from lung cancer was determined by the stage at which they were diagnosed. Although it is possible that someone may die from undiagnosed lung cancer (and the rate at which this happens is specifically estimated in the Natural History Model), this is very unlikely.

Other cause mortality

Other cause mortality was also estimated using data from NLST, primarily to provide consistency with the modelling of the natural history component and to reflect mortality in a population eligible for lung cancer screening. A Weibull model with proportional hazards for age, $PLCO_{m2012}$ linear predictor, current smoking status, and the interaction between $PLCO_{m2012}$ and current smoking status was found to be the best fit for overall mortality. Due to data from the NLST leading to estimated survival longer than would be expected in a cohort of current and former smokers within the general population, data from UK life tables were also modelled(33). This was achieved using a Gompertz model. As mortality in a cohort of current and former smokers would be expected to be greater than that for the general population, the hazard rates from the Weibull model (NLST) and Gompertz model (UK lifetables) were summed to produce survival curves that were more consistent with the expected survival of this cohort in the absence of lung cancer.

Effectiveness evidence

Risk prediction

In contrast to the original ENaBL model, it was assumed that the $PLCO_{m2012}$ risk prediction model(12) would be used to identify those individuals at high risk of lung cancer, and therefore more likely to benefit from a lung cancer screening programme. In the original model(5) the Liverpool Lung Project prediction model(LLPv2) (34) was used. However, the new natural history model incorporates risk as determined from the $PLCO_{m2012}$ (as used in NLST). Thus, for consistency, and to avoid the need for jointly modelling two risk scores, the $PLCO_{m2012}$ model was assumed to estimate risk of lung cancer in the modelled screening programmes. As the original use of LLPv2 did not constitute an endorsement of LLPv2, use of $PLCO_{m2012}$ in this updated model is not an endorsement, and the change is solely due to data availability in NLST.

The $PLCO_{m2012}$ risk prediction model includes the following: age, race or ethnic group, level of education, body mass index (BMI), chronic obstructive pulmonary disease (COPD), personal history of cancer, family history of lung cancer, smoker statues (current or former),

smoking intensity (average number of cigarettes per day), duration of smoking, smoking quit time(12).

The distribution of $PLCO_{m2012}$ in the population of all UK smokers was estimated by imputing $PLCO_{m2012}$ for individuals in the Health Survey for England 2019(35). Once the $PLCO_{m2012}$ was imputed for all individuals in the HSE dataset, statistical models were fitted to predict the risk score based on age and smoking status (current or ex-smoker). This was used to estimate risk of lung cancer for those simulated individuals in the DES model.

Although the HSE 2019 data included information on smoking history, a number of parameters in the $PLCO_{m2012}$ required imputation as they were not available in the HSE dataset. These were family history of lung cancer, personal history of cancer, personal history of COPD and cigarettes smoked per day for former smokers. Brief details on how this was achieved are given below.

Family history of lung cancer was imputed based on how many parents smoked regularly. This was estimated using HSE 2019 data on the probability that one or both parents smoked regularly, and the probability that an individual had a parent with a history of lung cancer (as reported in Lebrett 2020(36)). These analyses resulted in an estimated 16.5% of never-smokers having a positive family history of lung cancer. For comparison, Warkentin 2019(37) studied the UK Biobank and found that 11.7% of never-smokers had a positive family history of lung cancer.

Personal history of cancer was informed by smoking status, age and sex. Using the prevalence of all cancers in England from the National Cancer Registration and Analysis Service (NCRAS)(38), estimates on the percentage of cancers for which smoking was responsible, as reported in Brown 2018(39) were applied, and the relative risks for all cancers due to smoking were estimated. Similar figures were assumed to apply to prevalent cancers.

History of COPD was imputed using data on age and smoking duration. This was based on a logistic regression of 2010 Health Survey for England data(40) where the probability of an individual having been told by a doctor that they have COPD, bronchitis or emphysema was found to be related to age and duration of smoking. The coefficients for age and smoking duration were then applied to age and smoking duration data from 2019 HSE to impute personal history of COPD.

As number of cigarettes per day was not recorded in HSE 2019 for former smokers, it was estimated using historical trend data from the ONS(41). It was assumed that former smokers followed the general pattern of smoking in the year that they quit. Data from the

ONS on the average daily cigarettes consumed over the past decades were used alongside HSE 2019 data on years since quitting to approximate cigarettes smoked per day for former smokers.

Additional parameters also required some assumptions. Data on age from HSE 2019 were provided by subgroups (e.g. 16-17 years, 20-24 years), so assumptions were made to approximate actual age. Level of education was mapped to the highest educational qualification question from HSE 2019. Ethnicity was also mapped, with Black and Asian from HSE 2019 mapped to Black and Asian from $PLCO_{m2012}$ and everyone else mapped to White.

Once all relevant parameters were imputed, the $PLCO_{m2012}$ was calculated for each individual in the HSE 2019 dataset. A linear regression was applied with age and smoking status as independent variables. Since the $PLCO_{m2012}$ values were not normally distributed (nor did they have normally distributed residuals after adjusting for age and smoking status), skewed distributions were used. For current smokers a skewed T-distribution was used, whereas for former and never-smokers a skewed normal distribution was used. The resulting coefficient estimates and 95% CIs are shown in Table 10.

Table 10 Coefficient estimates (95% CIs) for $PLCO_{m2012}$ risk prediction based on HSE2019 data

Variable	Coefficient (95% CI)	
	Current smokers	Former smokers
Age*	0.119793 (0.116006, 0.12358)	0.087278 (0.08465, 0.089905)
Constant	-2.69352 (-2.80331,-2.58372)	-4.90745 (-5.07695, -4.73795)
Alpha	-4.28561 (-5.33874, -3.23247)	-0.95644 (-1.26978, -0.64309)
Omega	1.881782 (1.683074, 2.10395)	1.19585 (1.099787, 1.300305)
DF	2.291878 (1.917487, 2.739371)	NR

*Age was centred on 60 years; CI, confidence interval; DF, degrees of freedom parameter; NR, variable not relevant to skew normal linear regression

Uptake of LDCT screening

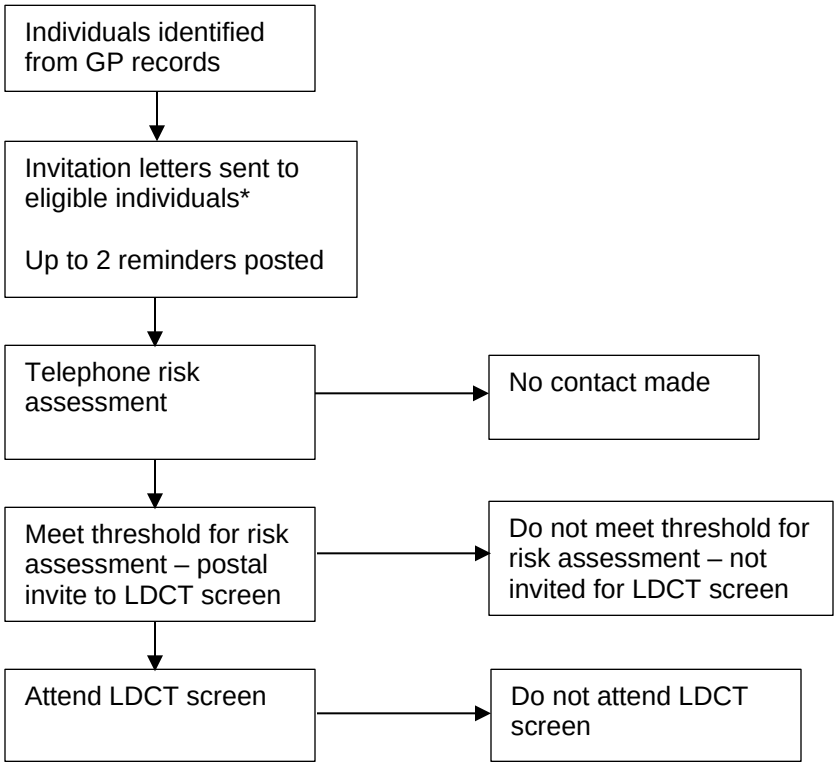
In the original report(5), estimates of screening uptake were taken from UKLS(3). Since the publication of Snowsill 2018(5), regional LDCT lung cancer screening programmes/pilots and trials have been conducted in the UK (Manchester LHC pilot(13), Liverpool Healthy Lung Programme (HLP)(42), West London lung cancer screening (LCS) pilot(43), Lung Screening Uptake Trial (LSUT)(44)). These have generally taken one of two approaches to

inviting individuals to be assessed for their eligibility to partake in LDCT screening for lung cancer. In both approaches, individuals are identified from GP records and then sent information in the post. They are then invited to either an in-person LHC to assess their eligibility for LDCT screening, or have this assessment via a telephone call.
To update the model we used recommendations from the Lung Cancer Screening Pathways Task and Finish Group(45) on the most appropriate pathway to inviting individuals for LDCT screening. The consensus from this group was that a telephone call to assess eligibility would be the most likely way of inviting individuals to take part in screening, should a national programme be commissioned. Thus, in our base case analysis we assumed that:

- Potentially eligible individuals were identified from GP records as being 55-80 years old and ever-smokers (including former and current smokers).
- Those identified were sent a letter inviting them to call for a telephone lung cancer risk assessment to evaluate their eligibility to attend a LDCT scan.
- Up to two reminder letters would be sent to individuals to make this call.
- During the telephone call, the risk of lung cancer for that individual was assessed.
- Those found to have a risk above the threshold would be invited to a LDCT scan, with an appointment letter sent to them.

The flow of individuals through the uptake pathway is shown in Figure 9. There are currently no published uptake results in studies where eligibility for screening is assessed via the telephone. The Yorkshire Lung Screening Trial (YLST)(46) and the SUMMIT study(47) are both using this pathway to identify eligible individuals. The aim of the YLST is to estimate participation in community-based screening, compare risk models for predicting those at high risk for lung cancer, and evaluate clinically-relevant outcomes from LDCT screening(46). We use data from the YLST, kindly provided by Mat Callister and his team (personal communication). Estimates of uptake, and the proportion of reminders sent are presented in Table 11, the resource use and costs associated with this pathway, were also taken from the YLST, and are presented in Table 16.

Figure 9 Assumed pathway to LDCT screening



*eligibility defined by smoking history and age.

Scenario analyses were conducted assuming uptake rates from the Manchester LHC pilot(13) (see Table 11). The proportion accepting an invite for risk assessment is much lower in the Manchester LHC than the YLST (28.5% vs 50.8%). This could be explained by the fact that in the Manchester LHC participants had to attend a face-to-face appointment for a risk assessment, while in the YLST the risk assessment was conducted over the telephone. In the Manchester LHC pilot, if individuals were eligible for a LDCT scan, they were offered it on that, or the following, day. While in the YLST, participants made an appointment over the telephone, if eligible, for a future LDCT. The additional costs associated with face-to-face risk assessment in the Manchester LHC are included in the scenario analysis.

It is further assumed that once a participant has entered the screening programme (by attending the first screen), they will participant in any further screening rounds they are invited to attend. Thus, screening compliance is assumed to be 100% for those entering the screening programme. Compared to data available in the published literature, this assumption is likely an over-estimate of screening compliance. Crosbie 2019(48) report that 90.2% of eligible participants returned for the second round of the Manchester LHC pilot,

while Horeweg 2014 report that in NELSON 97.8% of those eligible for their second screen attended the second screen and 96.5% of those eligible for the third screen attended their third screen(49).

Table 11. Screening uptake parameters used in the model

Parameter	**Base case analysis**		**Scenario analyses**	
	Mean (SE)	Source	Value	Source
Proportion accepting invite for eligibility risk assessment	50.8% (0.2%)	YLST (personal communication)	28.5%	Manchester LHC pilot(13)
Proportion accepting invite for LDCT scan	83.6% (0.6%)	YLST (personal communication)	96.9%	Manchester LHC pilot(13)
Proportion of those approached sent 1st reminders	78.8% (0.2%)	YLST (personal communication)		
Proportion of those approached sent 2nd reminders	52.7% (0.2%)	YLST (personal communication)		
Screening compliance	100%	Assumption		

LHC, Lung Health Check; YLST, Yorkshire Lung Screening Trial

LDCT specificity

In the original model, calculation of the specificity of LDCT from UKLS was incorrect. It has now been updated, again based on UKLS(3, 50), assuming that of 1942 participants in the screening arm who did not have a diagnosis of lung cancer within 12 months of LDCT screening, 72 were referred to the MDT for further investigation on the basis of their LDCT scan. Thus, a specificity of (1942-72)/1942 = 0.963 (95%CI 0.953, 0.970) was assumed. In the model, the specificity of LDCT linked to the costs and disutility associated with individuals who have a false positive LDCT scan. A false positive LDCT scan was defined by an individual being referred to the MDT for further investigation, but not receiving a diagnosis of lung cancer.

LDCT sensitivity

As described in the previous Chapter, LDCT sensitivity is estimated by stage from the natural history model which is calibrated to NLST. This also incorporated a final screen threshold effect to allow for the possibility that, for the final LDCT screen, there is a lower threshold on what constitutes a positive screen, due to this being the last time the individual is screened. This is a change to the original model where sensitivity was not dependent on stage at diagnosis, and did not include a final screen threshold effect.

Indeterminate LDCT screening results

Results from a LDCT screen will fall into one of three categories: no follow-up required in this screening round, immediate referral, or an indeterminate result. Those individuals with an indeterminate result require a follow-up LDCT scan or scans. If the follow-up LDCT(s) indicates a positive LDCT, individuals will be referred to the MDT for further investigation. If, however, the follow-up LDCT scan(s) are negative, the individual will re-join any further screening rounds. Based on the UKLS, it is assumed follow-up of an indeterminate LDCT scan would involve a repeat LDCT scan at 3 and at 12 months later, or only at 12 month later(3).

In the original model(5), data on the number of additional LDCT scans were taken from the UKLS where 47% of the sample were defined as having an indeterminate LDCT scan(3): 23% having a 3 month and 12 month scan, and 24% having a 12 month follow-up scan (Table 12). In 2015 the British Thoracic Society (BTS) updated their nodule management guidelines(51). To reflect these updates in the current model, data from UKLS(3) were used but as individuals would have gone through the new BTS guidelines (personal communication from David Baldwin, see Table 12).

Table 12 Proportion of participants having an indeterminant LDCT screening result

Source	% referred immediately	% having 3 and 12 month follow-up LDCT scan	% having 12 month follow-up LDCT scan only	% having no follow-up	Source
UKLS	3.2% 64/1994	23.7% 472/1994	24% 479/1994	49.1% 979/1994	Field(3)
UKLS (as applied to 2015 BTS nodule management guidance)	6%	12%	2%	80%	Field(3) and personal communication [David Baldwin]
Liverpool HLP	9%			81%	Ghimire (42)
Manchester LHC pilot, 1st round	4.7%	12.7%	Only 3 month follow-up scans given	82.6%	Crosbie (13)
Manchester LHC pilot, 2nd round	2%	6%	Only 3 month follow-up scans given	92%	Crosbie (48)
West London LCS pilot	1.7%	14.2%*		84.1%	Bartlett(43)

*14.2% represents the proportion of participants either having LDCT scans at 6 weeks, 3 months, 9 months and/or 12 months, or PET-CT scan. BTS, British Thoracic Society; HLP, Healthy Lung Programme; LCS, lung cancer screening; LHC, Lung Health Check; NA, not appropriate; UKLS, UK Lung Cancer Screening;

These updated estimates suggest 14% have indeterminate LDCT results, a much lower proportion than assumed in the original model(5). Published data from the Liverpool HLP(42), Manchester LHC pilot(13) and West London LCS pilot(43) are fairly consistent with this (see Table 12 above).

It was assumed that for the annual and triple LDCT screening strategy, individuals requiring a follow-up LDCT scan at 12 months would just re-enter the screening programme in the following year. However, for the final screen of the annual or triple screening strategies, individuals may be invited to attend a LDCT 12 months later. Thus, for the base case analysis the proportion of individuals having an indeterminate LDCT screening result, and therefore incurring the additional costs of one or more follow-up LDCT scans, were as given in Table 13.

Table 13 Assumptions on the proportion of individuals receiving a LDCT screen who have an indeterminant LDCT scan result for each modelled screening frequency

Frequency of LDCT screening strategy	% having 3 and 12 month follow-up LDCT scan	% having 12 month only follow-up LDCT scan
Base case		
Single (one-off)	12%	2%
Annual		NA – re-enter screening
Biennial		2%
Triple – 1st and 2nd rounds		NA – re-enter screening
Triple – 3rd round		2%
Scenario analysis		
Single	0%	14%
Annual		NA – re-enter screening
Biennial		14%
Triple – 1st and 2nd round		NA – re-enter screening
Triple – 3rd round		14%

Based on data from UKLS(3) and personal communication (David Baldwin). NA, not applicable.

Impact on survival

As with the original model, the modelling approach taken leads to a reduction in lung cancer mortality in terms of a stage shift at diagnosis, with those diagnosed via screening likely to be diagnosed at an earlier stage than if clinically presented. However, due to updates in the estimation of survival after a lung cancer diagnosis, survival may be improved even if a stage-shift is not observed. This is because screen-detected lung cancers are estimated to have better survival than lung cancers detected outside of screening, even if diagnosed in the same stage. See above section on **Mortality after a lung cancer diagnosis**.

Health-related preference-based outcomes

EQ-5D is the preferred method to measure health-related quality of life (HRQoL), with the UK time trade-off value set from a sample of the general population the preferred valuation. Due to changes in the natural history model, dis-utilities were sought for NSCLC stages IA1, IA2, IA3, IB, II, III, IV, SCLC limited and extensive stages, and screen-related events. The database search used in the original report was updated. Inclusion criteria were as in the previous report: primary studies using EQ-5D to measure HRQoL in patients with lung cancer (unless those patients were experiencing specific adverse events or symptoms), and systematic reviews of EQ-5D in patients with lung cancer. One reviewer screened all titles, abstracts and subsequent full-text articles. We identified 1,063 hits between January 2017 (the date of the previous database searches) and February 2021. An update search was conducted in July 2022, identifying a further 475 hits. Six studies reporting on utility by lung cancer stage and two studies reporting on screening-related utilities were subsequently screened at full-text.

Lung cancer stage

Two systematic reviews of HRQoL by lung cancer stage were identified. In Blom 2020(52), utilities from 27 studies are meta-analysed regardless of the method used to value the utility, and are reported for lung cancer overall or combined in stages I-II and stages III-IV (Table 6). Pourrahmat 2021(53) identified 5 studies reporting utilities by lung cancer stage. Among the identified studies in Blom 2020 and Pourrahmat 2021, none were conducted in the UK using the EQ-5D. The most appropriate study included in both reviews was Tramontano 2015(10) which was used in the original report to inform the lung cancer stages.

The study by Yang 2019(54) is set in Taiwan and includes 1715 patients with NSCLC lung cancer. EQ-5D values are reported using the Taiwan and the UK tariff, however, the authors report some adjustment of utility values from the Taiwan tariff to limit values between 0 and 1. It is not clear whether the same constraints have been placed on the UK tariff EQ-5D values. Due to this uncertainty and EQ-5D mean estimates being split by age and squamous cell carcinoma (for NSCLC), these values are difficult to apply to our model (Table 6). Zeng 2020(55) and Liu 2022(56) measured EQ-5D using the Chinese tariff and report mean values for stages I, II, III and IV (Table 6). The estimates from Liu are based on 347 individuals with NSCLC and SCLC lung cancer from 7 hospitals in China. Zeng 2020 based their estimates on 93 patients with lung cancer. Neither Liu 2022 or Zeng 2020 report utilities by stage for NSCLC and SCLC separately.

Kuehne 2022(57) report EQ-5D-5L values at lung cancer diagnosis using the Canadian tariff from 58 individuals with limited SCLC and 97 individuals with extensive SCLC within a cohort study from a single hospital. No other studies identified from the database searches reported utility values for individuals with SCLC by stage.

Given the lack of evidence identified in the update searches for EQ-5D utilities by NSCLC lung cancer stage, there was no reason to change the evidence source used in the original model, Tramontano 2015(10). This US study reported EQ-5D values for stages I, II, III and IV from 2396 individuals with lung cancer (Table 14). EQ-5D estimates from Kuehne 2022 were used to inform utility associated with SCLC.

Table 14 Published utility values for lung cancer by stage

Study	Participant characteristics	Valuation method, Tariff	Stage/category	Mean utility (uncertainty)
Systematic reviews				
Blom 2020(52)	Multiple countries, NSCLC and SCLC	Any included	All stages	0.68 (95%CI 0.61, 0.75)
			I-II	0.78 (95%CI 0.70, 0.86)
			III-IV	0.69 (95%CI 0.65, 0.73)
Pourrahmat 2022(53)	Multiple countries, NSCLC and SCLC	Any included	I	Range: 0.59 - 0.86
			II	Range: 0.56 – 0.81
			III	Range: 0.27 – 0.89
			IV	Range: 0.66 – 0.84
Individual studies				
Liu 2022(56)	China, NSCLC and SCLC, N=347	EQ-5D, Chinese	I	0.886 (SD 0.144)
			II	0.889 (SD 0.181)
			III	0.842 (SD 0.224)
			IV	0.819 (SD 0.218)
Zeng 2020(55)	China, Unclear if NSCLC and SCLC, N=93	EQ-5D, Chinese	I	0.8 to 0.9*
			II	~0.7
			III	~0.6
			IV	0.38 to 0.57*
Yang 2019(54)	Taiwan, NSCLC, N=1715	EQ-5D, UK	SqC I-IIIA (<65yrs)	0.80 (SE 0.03)
			SqC IIIB-IV (<65 yrs)	0.74 (SE 0.04)
			NSqC I-IIIA (<65 yrs)	0.84 (SE 0.01)
			NSqC IIIB-IV (<65 yrs)	0.77 (SE 0.01)
			SqC I-IIIA (≥65yrs)	0.78 (SE 0.02)
			SqC IIIB-IV (≥65 yrs)	0.61 (SE 0.04)
			NSqC I-IIIA (≥65 yrs)	0.79 (SE 0.01)
			NSqC IIIB-IV (≥65 yrs)	0.72 (SE 0.02)
Tramontano 2015(10)	US, NSCLC and SCLC N=2396	EQ-5D, US	I	0.81 (SD 0.17)
			II	0.77 (SD 0.17)
			III	0.77 (SD 0.18)
			IV	0.76 (SD 0.19)
Kuehne 2022(57)	Canada, SCLC, N=155	EQ-5D, Canadian	Limited	0.802 (0.77, 0.84)
			Extensive	0.718 (0.68, 0.76)

* depending on whether used EQ-5D-3L or EQ-5D-5L. CI, confidence interbal; NSCLC, non-small cell lung cancer; NSqC, non-squamous cell; SCLC, small cell lung cancer; SD, standard deviation; SE, standard error; SqC, squamous cell

Although there are many differences in the studies reporting EQ-5D utilities, the values assumed from Tramontano 2015(10) are somewhat consistent with those in Blom 2020(52) (although Tramontano 2015 was the largest study in that meta-analysis) and Yang 2020(54). While estimates from Zeng 2020(55) cover a much greater range than those assumed from Tramontano 2015: 0.9 – 0.4 (stage I – stage IV, for EQ-5D-5L), and 0.8 - 0.6 (stage I - stage IV, for EQ-5D-3L).

As in the original model, we assigned the utility for the stage at diagnosis for the remainder of the participants life, this was a simplifying assumption. The utilities reported by Tramontano 2015(10) were obtained within 6 months of participants receiving their lung cancer diagnosis. Additional analyses by Tramontano 2015 for participants followed up approximately 1 year later show a statistically significant reduction in EQ-5D utility for late stages. The addition of a Recurrence event to the DES model (see Section X), means that all simulated individuals are assumed to enter stage IV before dying from lung cancer (regardless of the stage at diagnosis). Thus, for the period of time individuals are assumed to be in Recurrence, they are assigned the disutility associated with stage IV NSCLC or extensive stage SCLC, reflecting a reduction in utility as lung cancer progresses. In a scenario analysis, the impact of assigning a lower utility to stage IV is evaluated (see Table 15).

Assignment of pre-clinical stage utilities
We assume that for pre-clinical lung cancer stages there is no decrement in utility. Individuals with pre-clinical cancer are therefore assigned the same utility as someone without cancer. Once a cancer is diagnosed, the individual is then assigned a stage-related disutility as described above. The only exception to this, is for individuals with pre-clinical stage IV lung cancer, who are assumed to have some disutility even though the cancer has not yet been diagnosed. This disutility is equivalent to that assumed for clinical stage IV.

LDCT screening
We specifically sought utilities related to LDCT screening, and to having a false positive LDCT result for lung cancer, which would include any unnecessary further investigations. To inform utilities related to LDCT screening, two studies were potentially relevant from the update searches: a primary lung cancer screening study in Canada(58), and a systematic review of disutilities for cancer screening(59). The systematic review by Li 2019(59) did not include any new evidence specific to lung cancer screening from that already reviewed in Snowsill 2018(5).

The primary study by Taghizadeh 2019(58) reported EQ-5D from 1237 individuals undergoing LDCT screening using the Canadian tariff. Taghizadeh 2019(58) reported no difference in EQ-5D values at study enrolment, or 1 month and 12 months after receiving the LDCT scan result. For individuals who received a positive LDCT scan result, no differences in EQ-5D values were observed at study enrolment, compared to values 1 month after their positive result, 1 month after additional follow-up LDCT or other tests, and 12 months after enrolment. Separate analyses for those participants found to have false positive LDCT scan results are not reported. As well as measuring EQ-5D, Taghizadeh 2019 measured anxiety using the State Trait Anxiety Inventory (STAI)-State anxiety score.

The only statistically significant change from baseline reported in Taghizadeh 2019 was for the STAI-State anxiety score at 1 month for all individuals.

Taghizadeh 2019(58) is the only study we are aware of that measures EQ-5D associated with LDCT screening. Based on the EQ-5D data reported in this study, suggesting no disutility associated with a LDCT scan, nor with a positive LDCT scan (including any further investigations), the mean disutilities in the base case analysis were assumed to be 0.00 (see Table 9).

However, in a scenario analysis, to reflect evidence of a change in anxiety score as seen on the anxiety-specific questionnaire used by Taghizadeh 2019(58) ((STAI)-State anxiety score), and other evidence as discussed in the previous report(5), disutilities for a LDCT scan and false positive result were assumed:

- A disutility of 0.01 for a 2 week period associated with a LDCT scan
- A disutility of 0.063 for a 3 month period associated with a false positive LDCT scan result.

The utilities and disutilities assumed in base case and scenario analyses are shown in Table 15.

Individuals eligible for lung cancer screening

Due to the targeted nature of the modelled screening programmes, individuals eligible for lung cancer screening were current and former smokers, and so were unlikely to have the same health-related quality of life as seen for the general population. Thus, as in the original model, the health-related quality of life of the simulated population was estimated using data from Health Survey England 2014(60). See Snowsill 2018 for further details(5). This analysis provided EQ-5D utility values for current and former smokers (in the absence of lung cancer) by sex: 0.746 for females and 0.774 for males.

Implementation of utilities

As seen in the above sections, the published utility values associated with individuals having a diagnosis of lung cancer(10, 57) exceed those estimated for the population without lung cancer. Therefore, to avoid a lack of face validity with the utilities implemented in the model, it was assumed that there is no disutility associated with stage I NSCLC and limited SCLC. For later stages, stage specific disutilities, informed by these published sources, were applied. Table 15 details the utilities for the eligible screening population, and the disutilities associated with lung cancer and LDCT screening that were applied in the model.

Table 15 Utilities and disutilities implemented in base case and scenario analyses

Event	Base case analysis		Scenario analyses	
	Mean (SE)	**Source**	**Mean**	**Source**
Utilities				
Males eligible for lung cancer screening	0.820	**HSE 2014**		
Females eligible for lung cancer screening	0.791	**HSE 2014**		
Disutilities vs no lung cancer				
NSCLC: IA1, IA2, IA3 and IB	0	Tramontano(10)		
NSCLC: II	0.04 (0.013)			
NSCLC: III	0.04 (0.009)			
NSCLC: IV	0.05 (0.010)		0.252	Sturza(61)
SCLC: limited	0	Kuehne 2022		
SCLC: extensive	0.08 (0.027)			
LDCT screen	0.00 (0.008)	Taghizadeh(58)	0.01 (for 2 weeks)	NELSON(62)
False positive result	0.00 (0.015)		0.063 (for 3 months)	Mazzone(63)

NELSON, Nederlands–Leuvens Longkanker Screenings Onderzoek; NSCLC, non-small cell lung cancer: SCLC, small cell lung cancer; SE, standard error

Resource use and costs

Programme administration costs

Two information technology (IT) costs were included in the cost analysis since they are variable and distinct from fixed or transactional costs related to setting in place a programme. The current annual cost of extraction of executable data from primary care IT systems is estimated as £36,000 based on a government contract price for an existing centralised patient index service(64). This was applied as a one-off cost for the identification of all age-eligible ever-smoker individuals for potential participation because in this closed population model we do not consider people ageing into eligibility (or having their risk re-estimated after some time). Spread over an estimated 13 million ever-smokers in England(65), the per person cost (for the purposes of cost-effectiveness analysis) is under 1p and is therefore negligible. However, the cost of creating and running a dedicated digital database of participants is enduring. This is estimated as £2.06 per subject per year, based on databasing of Abdominal Aortic Aneurysm screening(66). It was assumed that this cost was no longer applicable after death or after the scheduled end of the programme.

The costs per item of resource required for the intended approach to programme recruitment, based on written correspondence and telephone triaging, were obtained from audited costs accrued by the YLST (thanks to Professor Matthew Callister, Leeds Teaching Hospital NHS Trust). Pre-invitation letters/notices were sent to all ever-smokers within the entry age criteria of each screening programme design; invitation to participate letters were sent as follow-up to the same group – with reminders to those failing to respond; triage telephone calls - including attempts - were made to the positive subgroup thereof; and screen appointment letters were sent to the further subgroup identified as meeting the programme risk criteria. See Table 16.

Table 16 Unit costs of resources for programme recruitment

Resource	Unit cost (£)
Pre-invitation notice	£0.66
Invitation letter for participation	£0.79
Invitation reminder (as needed*)	£0.84
Telephone triage call	£7.62
Invitation letter for screening appointment	£0.70

*72% received a first reminder; 55% received a second reminder (based on the YLST experience)

LDCT

On the day of the screening the included resources were the LDCT scan itself and the cost of nurse time in support. This was approximated as 15 minutes of band 4 hospital nurse equivalent (unit cost £7.75). LDCT costs was assumed to be the same, whether it be the single one-off screen of the single screening programme design, or the 8th annual screen of the annual screening programme design. The unit cost of this LDCT in the base case was the weighted mean cost of all records in the NHS Reference cost schedule 2019/2020(67). Alternative unit costs were examined in scenario analyses. These were the cost of LDCT used in the original model(5), and the currency-adjusted cost of LDCT used in the recent Australian economic evaluation of LDCT screening in high risk groups (68). See Table 17.

Table 17 Unit costing of LDCT

LDCT setting	Unit cost (£)	Source
Base case		
Weighted estimate from all settings	£77.31	NHS Reference cost schedule 2019/2020: Direct access £88.31; Outpatient 1 £91.31; Outpatient 2 £72.47; Other £94.47 [HRG RD20A Imaging: Computerised Tomography Scan of One Area, without Contrast, 19 years and over]
Scenario analyses		
Weighted estimate from all settings	£98.80	NHS Reference cost schedule 2015/2016(69)
Australian cost	£143.44	Harpaz 2022(68)

Resource use for false positives

False positives are defined as those individuals referred to the MDT on the basis of LDCT screening results, but who do not receive a diagnosis of lung cancer within 12 months of the LDCT screen. Resource use for these individuals was informed by the UKLS(3) and unit costs were sourced from the NHS Reference cost schedule 2019/2020(67), see Table 18.

Table 18 Resource consumption and unit costs associated with false positive cases

Investigations/treatments received	Proportion of those referred to MDT but not found to have lung cancer (N=72 from UKLS)	Unit cost per intervention
MDT meeting	100%	£116.81
Further CT scan	84.7%	£77.31
Out-patient follow-up	30%	£151.13
PET scan	18%	£665.58
Needle biopsy	9.7%	£724.09
Surgical referral	5.5%	£57.00
EBUS	1.4%	£973.56
Oncology referral	1.4%	£59.50

CT, computed tomography; EBUS, endobronchial ultrasound; MDT, multidisciplinary team; PET, positron emission tomography; UKLS, UK Lung Cancer Screening

Indeterminate cases

These cases were defined by their requirement for further follow-up with subsequent LDCT at 3 or 12 months (see Table 13). For screening programmes with routine subsequent screens at 12 months, the source estimate (of requirement) was adjusted to zero at that timepoint.

Lung cancer costs

As originally, lung cancer care included resources for diagnostic imaging, surgery, radiotherapy, and medical therapeutic intervention. However, the approach to deriving estimates of uptake were revised according to an improved method of micro-costing. The 2014 Cancer Research UK list of lung cancer resources given in 'Saving Lives Averting Costs' (70) was adapted to reflect available resource options for patients diagnosed in 2018 and focussed on consumption by stage at diagnosis for good performers (PS 0-2) aged 55-77. The process of adaptation was by led by Dr David Baldwin with consensus from the Clinical Expert Group for lung cancer, hosted by the Roy Castle Lung Cancer Foundation (previously NHS England 2014 to 2020).

Consumption rates post-diagnosis were adjusted according to survival and applied at the time of event, both improvements of particular relevance to the high costs associated with recurrence. Separately, ongoing monitoring resources were applied for five years post-diagnosis or until death except for six months after diagnosis and recurrence, to avoid double counting. The lung cancer associated costs were derived for the main NSCLC stages: I, II, III and IV. To align with the NSCLC substages and SCLC stages modelled, it is assumed that all NSCLC stage I costs can be applied to NSCLC stages IA1, IA2, IA3 and IB; cost for limited SCLC are equivalent to NSCLC stage II costs, with extensive SCLC costs equivalent to NSCLC stage IV costs. See Table 19, Table 20 and Table 21 for the stage-specific costs.

Table 19 Rate of consumption and unit cost of lung cancer resources by stage at diagnosis* - Diagnostics, Surgery and Radiotherapy

Type	Intervention	Unit cost (£)	Diagnosis				Ongoing				Recurrence			
			Stage I	Stage II	Stage III	Stage IV	Stage I	Stage II	Stage III	Stage IV	Stage I	Stage II	Stage III	Stage IV
Diagnostics and imaging	Chest X-ray	£42	90%	90%	90%	90%	200%	200%	200%	200%	90%	90%	90%	80%
	Contrast enhanced chest, lower neck, and abdomen CT	£115	99%	99%	99%	95%	120%	120%	140%	200%	99%	99%	90%	60%
	PET-CT	£303	80%	80%	80%	15%	5%	5%	5%		30%	30%	5%	15%
	Spirometry	£146	85%	85%	80%	25%	10%	10%			30%	30%		
	T[L]CO test	£130	70%	70%	50%	10%					10%	10%		
	Flexible bronchoscopy alone - no EBUS	£652	10%	10%	10%	18%	10%	10%	10%	5%	10%	10%	10%	5%
	EBUS-guided TBNA plus or minus bronchoscopy	£749	20%	40%	80%	50%	10%	10%	10%	10%	50%	50%	15%	10%
	CT biopsy	£181	60%	40%	5%	30%					15%	30%	5%	5%
Surgery	Elective - Lobectomy, wedge resection, pneumonectomy, segmental resection, sleeve resection	£4,357	67%	65%	18%	2%								
	Emergency - Lobectomy, wedge resection, pneumonectomy, segmental resection, sleeve resection	£6,303	1%	1%	1%						1%	65%	18%	2%
	Airway stents for endobronchial obstruction	£1,515				5%						1%	1%	2%
	Endobronchial debulking	£7,720				5%								2%
Radio-therapy	Intracranial procedures	£3,084				1%								1%
	RT for curative intent (SABR)	£3,999	21%			2%								
	RT for curative intent (non-SABR)	£3,440	2%	23%	10%	2%					4%		10%	
	Palliative RT	£917			14%	60%							14%	60%

*Stage I costs are applied to all NSCLC stage I substages; Stage II costs are applied to NSCLC stage 2 and SCLC limited stage; Stage III costs are applied to NSCLC stage III; Stage IV costs are applied to NSCLC stage IV and SCLC extensive stage. CT, computed tomography; EBUS, endobronchial ultrasound; PET-CT, positron emission ultrasound – computed tomography; RT, radiotherapy; SABR, stereotactic ablative radiotherapy TBNA, transbronchial needle aspiration; T[L]CO, transfer factor for carbon monoxide;

Table 20 Rate of consumption and unit costs of lung cancer resources by stage* at diagnosis – therapeutic treatment

Type	Intervention	Unit cost (£)	Diagnosis				Recurrence			
			Stage I	Stage II	Stage III	Stage IV	Stage I	Stage II	Stage III	Stage IV
Chemo-radiotherapy / IO	Curative intent	£5,505			33%	0%	20%	20%		
Chemotherapy	Docetaxel monotherapy	£3,832				4%	4%			4%
	Docetaxel plus nintedanib	£8,743							5%	5%
	Pemetrexed maintenance	£12,576				5%	5%			4%
Chemotherapy doublet	platinum + vinorelbine (adjuvant in stage II)	£1,506		39%	7%	3%	3%			3%
	Gemcitabine + carboplatin	£2,372				4%	4%			4%
	Gemcitabine + cisplatin	£1,844				1%	1%			1%
	Pemetrexed + platinum	£5,869				4%	4%			4%
Immunotherapy	Pembrolizumab	£18,859			10%	14%	14%	10%	14%	14%
	Durvalumab	£32,699			20%			20%		
Chemo-immunotherapy	Pembrolizumab + carbo/cis + gem/pemetrexed	£23,901			15%	28%	28%	28%	28%	
TKI	Gefitinib	£13,788				1%	1%	1%	1%	
	Erlotinib	£3,063			1%	1%	1%	1%	1%	
	Entrectinib	£41,149				1%	1%	1%	1%	
	Crizotinib	£55,090				1%	1%	1%	1%	
	Alectinib	£87,630			1%	1%	1%	1%	1%	
	Osimertinib	£28,451				3%				
	Afatinib	£42,756			1%	1%	1%	1%	1%	1%
	Brigatinib	£55,290			33%	1%	1%	1%	1%	

*Stage I costs are applied to all NSCLC stage I substages; Stage II costs are applied to NSCLC stage 2 and SCLC limited stage; Stage III costs are applied to NSCLC stage III; Stage IV costs are applied to NSCLC stage IV and SCLC extensive stage. IO, immunotherapy; TKI, tyrosine kinase inhibitor

Table 21 Summary lung cancer resource costs by stage* at diagnosis and type

Cost type	Stage I			Stage II			Stage III			Stage IV		
	Diagnosis	Recurrence	On-going	Diagnosis	Recurrence	On-going	Diagnosis	Recurrence	Ongoing	Diagnosis	Recurrence	On-going
Diagnostics	£933	£766	£392	£1,046	£793	£392	£1,249	£343	£401	£788	£265	£422
Surgery	£2,982	£63		£2,895	£4,112		£847	£1,150		£549	£311	
RT	£909	£138		£791			£472	£472		£730	£581	
Chemo-rad		£1,101			£1,101		£1,817					
CT		£782						£437		£782	£656	
CT doublet		£393		£587			£105			£393	£393	
Immunotherapy		£2,640			£8,426		£8,426	£2,640		£2,640	£2,640	
Chemo-immunotherapy		£6,692			£6,692		£3,585	£6,692		£6,692		
TKI		£2,988			£2,988		£1,334	£2,988		£3,772	£428	
Follow-up	£271	£271		£217	£108		£162	£108		£108	£108	
Total	£5,094	£15,834	£392	£5,537	£24,220	£392	£17,999	£14,831	£401	£16,456	£5,382	£422

*Stage I costs are applied to all NSCLC stage I substages; Stage II costs are applied to NSCLC stage 2 and SCLC limited stage; Stage III costs are applied to NSCLC stage III; Stage IV costs are applied to NSCLC stage IV and SCLC extensive stage. CT, computed tomography; Chemo-rad, chemoradiotherapy; RT, radiotherapy; TKI, tyrosine kinase inhibitor;

Summary of changes made to original model

Table 22 Summary and justification of updates to original model(5)

Aspect of model	Original model(5)	Updated model	Justification
Structure			
Natural history model	Based on NLST data to estimate the risks of developing preclinical (occult) lung cancer, progression of preclinical lung cancer (through seven lung cancer stage IA, IB, IIA, IIB, IIIA, IIIB and IV), and the presentation (symptomatic or incidental) of lung cancer.	Completely new natural history model based on NLST data through seven NSCLC stages (IA1, IA2, IA3, IB, II, III and IV), and two SCLC stages (limited and extensive), the presentation (symptomatic or incidental) of lung cancer, incorporating additional heterogeneity, and allowing stage-specific estimates of LDCT sensitivity.	The original natural history model estimated stage distributions at lung cancer diagnosis in the presence or absence of screening that are not well matched to the stage distributions observed in LDCT trials and national statistics, with an overestimation of late stage cancers and underestimation of early stage cancers.
Recurrence	Clinical stage progression is implied since costs are informed by stage at diagnosis and utility is adjusted according to the stages remining to be lived.	Recurrence is added as an event, allowing a time-specific attribution of mean costs and disutilties at recurrence, including the relatively high cost of new TKI and immunotherapy options.	High cost interventions used at progression/recurrence have become a larger proportion of overall disease costs in recent years. Including a recurrence event allows these costs to be included after diagnosis – important in the context of discounting future costs. It also allows explicit accounting of disutilties at recurrence.
Parameters			
Screening uptake parameters	From UKLS(3)	From YLST (personal communication, Mat Callister)	Based on the Pathways Task and Finish group recommendations, likely pathway to screening programme to be through telephone triage (as in YLST) rather than through post questionnaire (as in UKLS).
Disutilities for screening events	Based on EQ-5D VAS results from NELSON(62) and EQ-5D results for chest x-ray screening(63)	Based on EQ-5D TTO values from LDCT screening study(58)	Updated data source more representative of research question, specifically EQ-5D TTO values for LDCT screening.

Aspect of model	Original model(5)	Updated model	Justification
Disutilities for pre-clinical lung cancer	Pre-clinical stages were assumed to attract the same stage utilities as clinical stages on the basis that diagnosis would not impact symptom based well-being. Stages attract an increasing disutility relative to stage IA/B. Tramontano 2015(10)	Pre-clinical stages IIA/B and IIIA/B do not attract a disutility relative to pre-clinical stages IA/B; pre-clinical stage IV retains the relative disutility.	To account for the likely psychological impact of diagnosis, and the physical impact of cancer treatment.
Proportion of screened individuals having follow-up LDCT scans for indeterminate results	Based on UKLS(3)	Updated based on UKLS(3) as applied to current BTS nodule management guidance(51)	Updated data more representative of likely screening programme (and consistent with UK LDCT screening pilots/programmes)
LDCT specificity	Based on incorrect data from UKLS	Updated calculation	Correction in calculation of specificity from UKLS. A false positive case is defined as someone referred to the MDT for further investigation but did not receive a diagnosis of lung cancer within the following 12 month period.
LDCT sensitivity	A single value estimated from the natural history model	Estimated from the natural history model by lung cancer stage at diagnosis, allowing for a final screen threshold effect	Likely to see greater sensitivity of LDCT as stage at diagnosis increase. Also, possible that the threshold for a positive screen is lowered for the final screen, so as not to miss any cancers.
Lung cancer mortality	Based on Goldstraw 2007(71)	Survival modelled by route of detection using data from NLST	Clinical opinion and empirical evidence indicate that LC survival depends on the route of detection. Data available from the NLST allows for calculation of LC survival by route of detection. In scenario analyses, LC mortality was estimated using international data from Goldstraw 2016 adjusted to reflect reduced survival in UK (using 1-year survival from NLCA)

Aspect of model	Original model(5)	Updated model	Justification
Other cause mortality	Based on adjusted lifetables	Estimated using a combination of NLST data and UK lifetables.	Primary reason for using NLST data was consistency with estimation of lung cancer survival and data used in the natural history model. However, as data from the NLST lead to estimated survival that was longer than would be expected in a cohort of current and former smokers within the general population, data from UK life tables were also modelled.
Impact of LDCT screening on survival	Based on a stage shift at diagnosis.	Survival modelled by route of detection using data from NLST	To reflect better survival often seen in screen-detected lung cancers vs interval or clinically presented lung cancers. Even if a cancer is screen detected in the same stage as it would have presented, there is now the possibility of improved survival.
Screening programme costs	Based on invitation and triaging by postal correspondence. Unit costs from ten Haaf 2017(6)	Based on telephone triaging and postal correspondence. Unit costs from YLCT (personal communication, Matthew Callister)	New UK regional evidence has become available to support a telephone triage approach (including uptake and joining rates), and provide health system specific unit cost estimates.
LDCT costs	Based on the unit cost of a direct access Computerised Tomography Scan of one area, without contrast, 19 years and over (HRG RD20A)(69)	Based on a weighted average unit price across all settings (same HRG), using updated NHS reference costs(67)	The preference of expert clinicians consulted as part of the revision of costs was for an 'all-settings' approach because it could not be assumed that the Direct Access setting better described the setting anticipated for LDCTs as part of a national screening programme.
Diagnosis and treatment costs	Based on a two-year costing approach, with index year costs from a UK teaching hospital(72), and second year costs estimated from the index year using a subsequent year ration from database analysis in England(73)	A five-year micro-costing approach which discerns costs at diagnosis from those at recurrence and through follow-up. Consumption is based on the most recent NLCA secondary care estimates for PS 0-2 in the 55-75 year age range.	Clinical expert feedback from workshops supporting the update of this research indicated a likely underestimation of late stage treatment costs due to the emergence of new technologies since the time of the original source estimate.

BTS, British Thoracic Society; HRG, healthcare resource group; NELSON, Nederlands–Leuvens Longkanker Screenings Onderzoek; NLCA, National Lung Cancer Audit; MDT, multidisciplinary team; PS, performance status; TKI, tyrosine kinase inhibitor; UKLS, UK Lung Cancer Screening; VAS, visual analogue scale; YLST, Yorkshire Lung Screening Trial;

Quality assurance

Quality assurance of the DES model was undertaken by the lead model developer, and all other members of the team. The quality assurance consisted of considerable checking of the model code and wiring, as well as inspecting primary and secondary outcomes for their face validity given the model inputs and assumptions.

Results

Naming convention

As in the original report(5), LDCT screening strategies are referred to in terms of: frequency – lower age limit – upper age limit – risk threshold.

Frequency is S for single, T for triple (annual scans for 3 years), A for annual and B for biennial. Lower age limit is either 55 or 60 years old, upper age limit is either 75 or 80 years old, and risk threshold is either 1.5(%),2.5(%) or 5(%).

The screening strategy S-55-75-1.5 therefore represents a single screen for individuals aged 55-75 years who have a predicted risk of lung cancer of ≥1.5%.

In the results, we distinguish between the eligible population for screening - those individuals aged 55-80 years old who have ever smoked would be invited to have a lung cancer risk assessment - and the individuals who join a screening programme – those meeting the age limits and risk thresholds for the different strategies and have agreed to participate. As will be seen further in the results section, the individuals joining the programmes (“joiners”) are a very small proportion of the eligible population. This is important as the benefits received by those joining a screening programme appear small when averaged out across the eligible population. Presenting results using “joiners” as the denominator as well as the whole population helps with understanding this dilution effect.

Base case analyses

Forty-eight hypothetical screening programmes were modelled, as well as a no-screening comparator arm, representing current practice. The main results focus on the probabilistic analysis, where parameter values were sampled from relevant distributions to reflect parameter uncertainty. Results from a microsimulation analysis, where all parameters were set at their central estimates, are also presented illustrating more detailed results, and allowing exploration of one-way sensitivity analyses and scenario analyses.

Probabilistic base case analysis

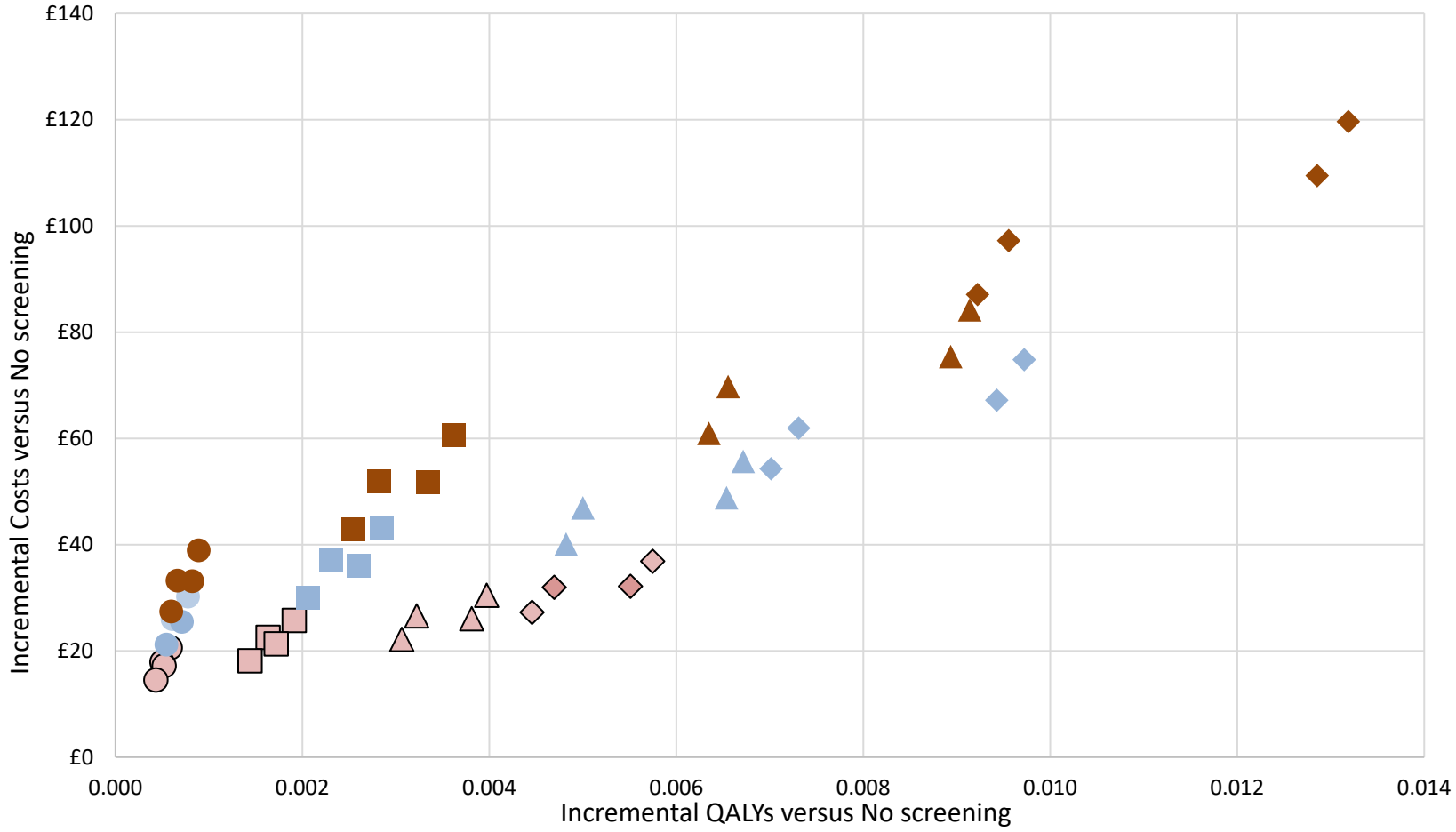
The probabilistic base case results presented below were conducted by simulating 500 cohorts of 3,000 individuals, thus represents 1.5 million independently simulated individuals. With cohort sizes of 3,000, it is likely that stability was not reached for strategy

mean costs and QALYs for each parameter value, and that Monte Carlo variability affects the apparent variability in the probabilistic results. Nevertheless, there should be adequate exploration of the parameter space with 500 samples and, with a total of 1.5 million simulations, the mean total costs and QALYs should be estimated with good precision.

Figure 10 shows the incremental QALYs and costs for all screening strategies compared to no screening. All screening strategies were associated with QALY gains and increased costs compared to the no screening strategy. The incremental QALYs ranged from 0.0004 (S-60-75-5%) to 0.0132 (A-55-80-1.5%), with incremental costs ranging from £15 (S-60-75-5%) to £120 (A-55-80-1.5%) per person aged 55-80 years over a lifetime. The incremental QALYs and costs were very small, and would not generally be considered significant (a QALY gain of 0.0132 is less than 5 quality-adjusted days). However, the gains from screening are concentrated in those individuals who join the screening programme, and are diagnosed with lung cancer before it would have presented clinically.

A pattern was observed where single screens generally lead to the smallest QALY gains and the smallest incremental costs. The triple screening strategy was estimated to lead to more QALY gains than the single screen, but at increased costs. The biennial screening strategies generally led to even more QALY gains and costs incurred than the triple screening strategies, with annual screening strategies estimated to produce the largest QALY gains, but with the highest costs. This pattern is seen in Figure 10.

Figure 10 Incremental QALYs and costs for all screening strategies compared to no screening

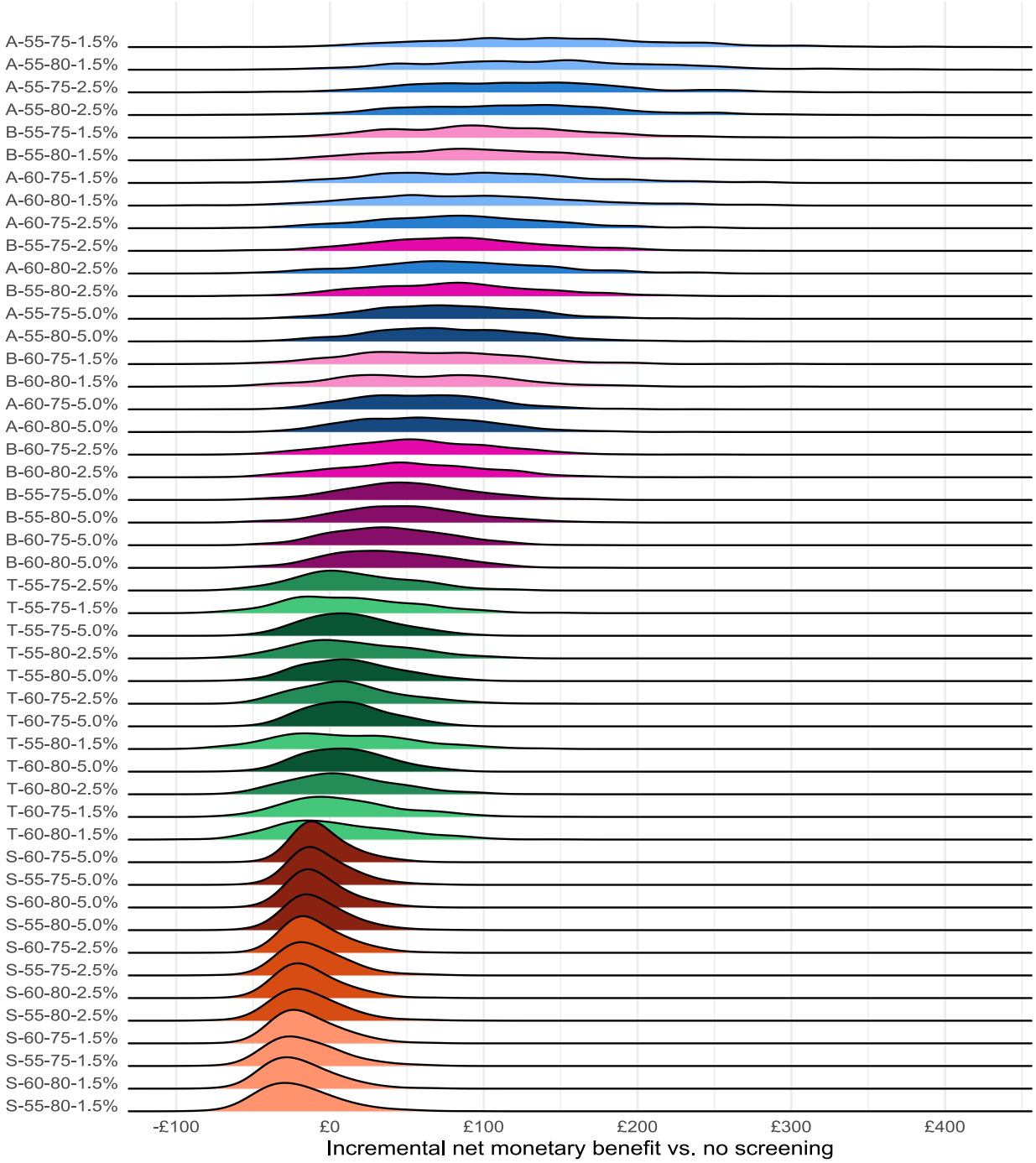


Legend: Diamonds represent annual screening strategies, triangles represent biennial screening strategies, squares represent triple screening strategies, circles represent the single screening strategies. The brown filled shapes represent those strategies with a risk threshold of 1.5%, blue filled shapes represent strategies with a 2.5% risk threshold; pink-filled shapes represent strategies with a 5% risk threshold.

Figure 11 below – known as a joy plot - shows the INMB at a willingness to pay threshold of £20,000 per QALY gained for each of the 48 screening strategies compared to no screening. Importantly, this plot allows for the uncertainty in the INMB estimation to be reflected. The strategies are ordered by mean incremental net monetary benefit, with the strategy with the smallest mean INMB (S-55-80-1.5%) at the bottom, and the strategy with the largest mean INMB (A-55-75-1.5%) at the top. Single screening strategies are shown in brown, triple screening strategies in green, biennial screening strategies in purple and single screening strategies in blue. The different shades within specific screening frequencies represent the lung cancer risk thresholds for the eligible populations, with the lightest shade reflecting the 1.5% risk threshold and the darkest shade reflecting the 5% risk threshold.

As can be seen in this figure, the single screening strategies do not look to be cost-effective at a willingness to pay of £20,000 per QALY gained: the mean INMBs are negative for all the single screen strategies. For the triple screening strategies, there is uncertainty as to whether they could be considered cost-effective, as their distributions cover a range of values either side of £0 INMB. For the biennial and annual screening strategies, the results indicate that these strategies could be considered cost-effective. Although there is much more uncertainty in the distributions for the these strategies than for the single and triple strategies, the vast majority of the distributions for annual and biennial strategies indicate positive INMBs. Furthermore, due to only 3000 individuals being simulated in the microsimulation, the probabilistic analyses represent Monte Carlo variation. Thus, should the microsimulation be run with >3000, the Monte Carlo variation would reduce, leading to the distributions illustrated in Figure 11 being associated with reduced uncertainty.

Figure 11 Joy plot showing the distribution of INMBs for each screening strategy compared to no screening at a willingness to pay of £20,000 per QALY gained



The following 4 figures (Figure 12, Figure 13, Figure 14 and Figure 15) represent the probability of each screening strategy being more cost-effective than no screening, showing the same pattern across strategies as seen in the joy plot (Figure 11 above). Looking across the figures, it can be seen that the annual and biennial strategies are associated with the highest probabilities of being more cost-effective than no screening (see Figure 12 and Figure 13). At a WTP of £20,000, all the annual and biennial screening strategies have >80% probability of being more cost-effective than no screening. For the triple screening strategies, most have a 50% probability of being more cost-effective than no screening at a willingness to pay of £20,000 (Figure 14). Even at a WTP of £50,000 per QALY gained, the single screening strategies have a maximum of 50% probability of being more cost-effective than no screening (Figure 15).

Figure 12 Probability of the annual screening strategies being more cost-effective than no screening

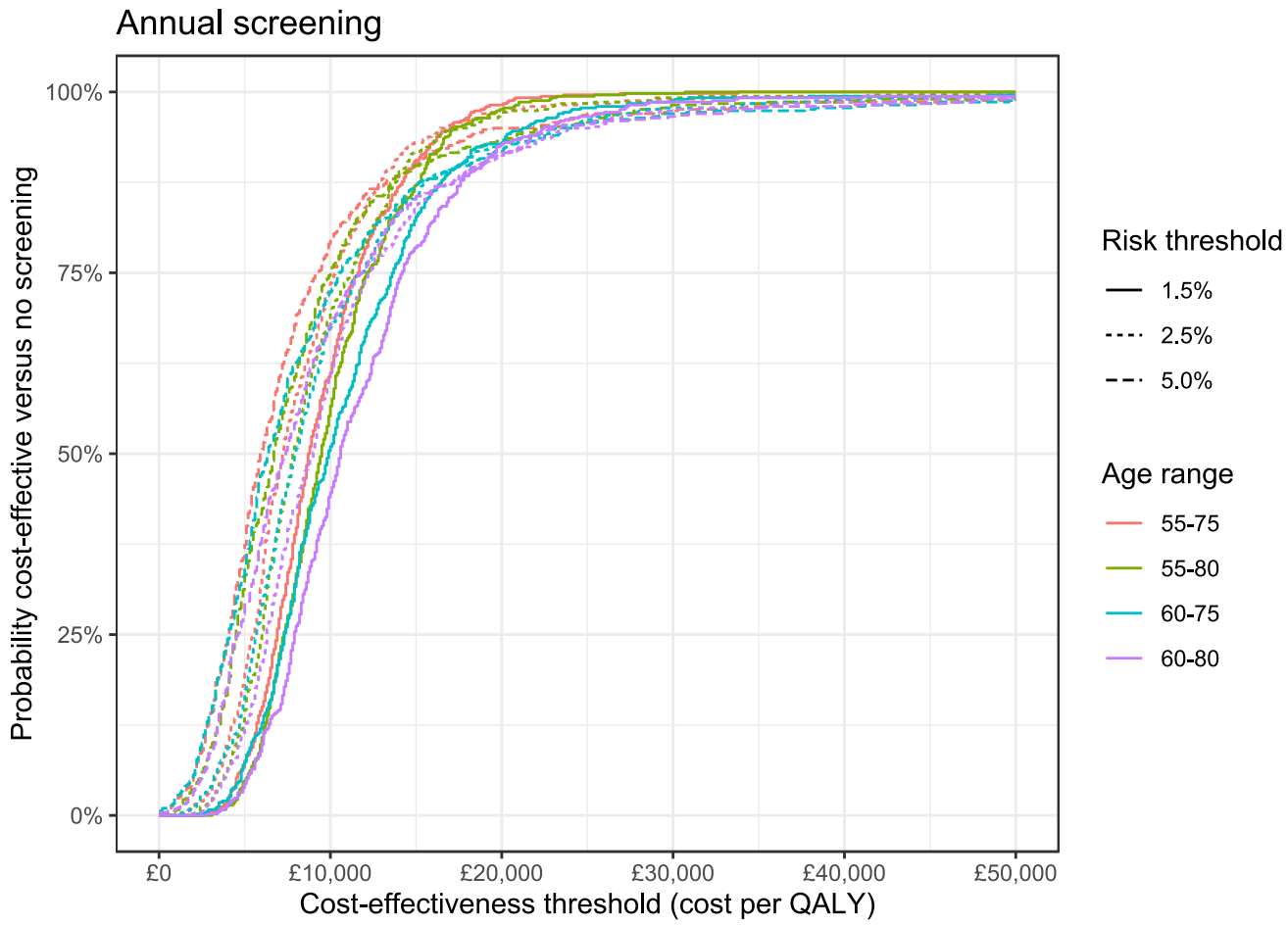


Figure 13 Probability of the biennial screening strategies being more cost-effective than no screening

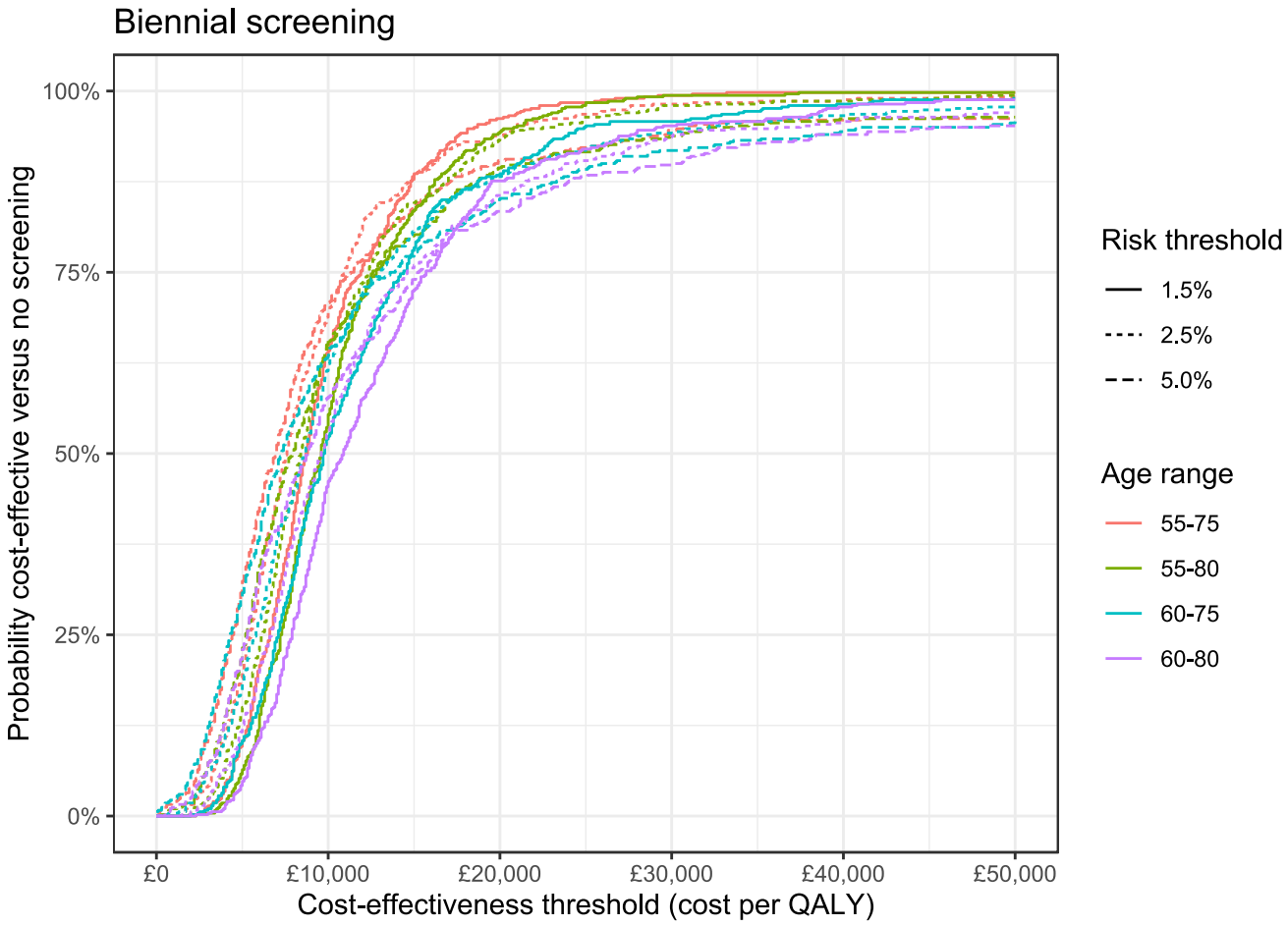


Figure 14 Probability of the triple screening strategies being more cost-effective than no screening

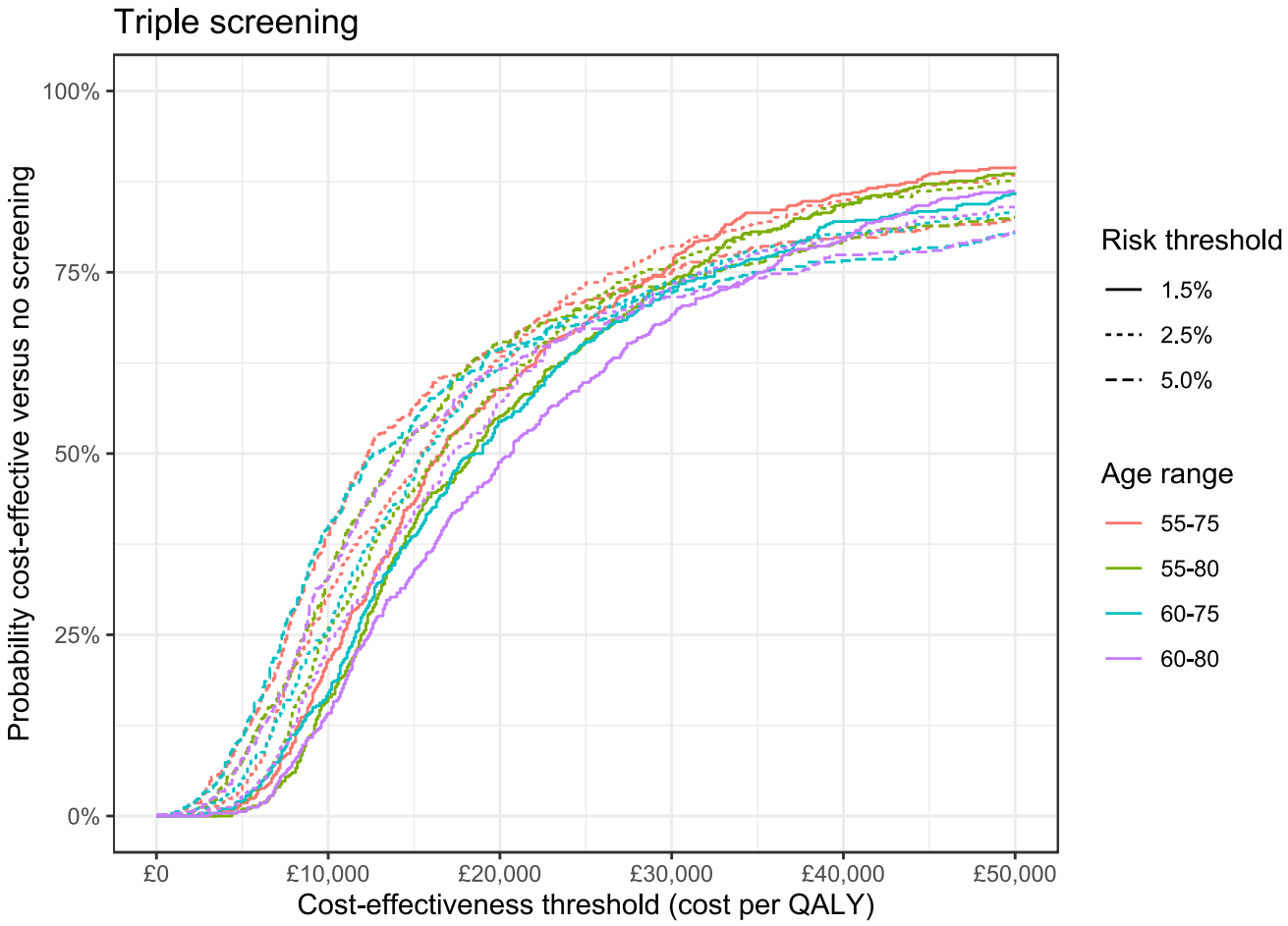
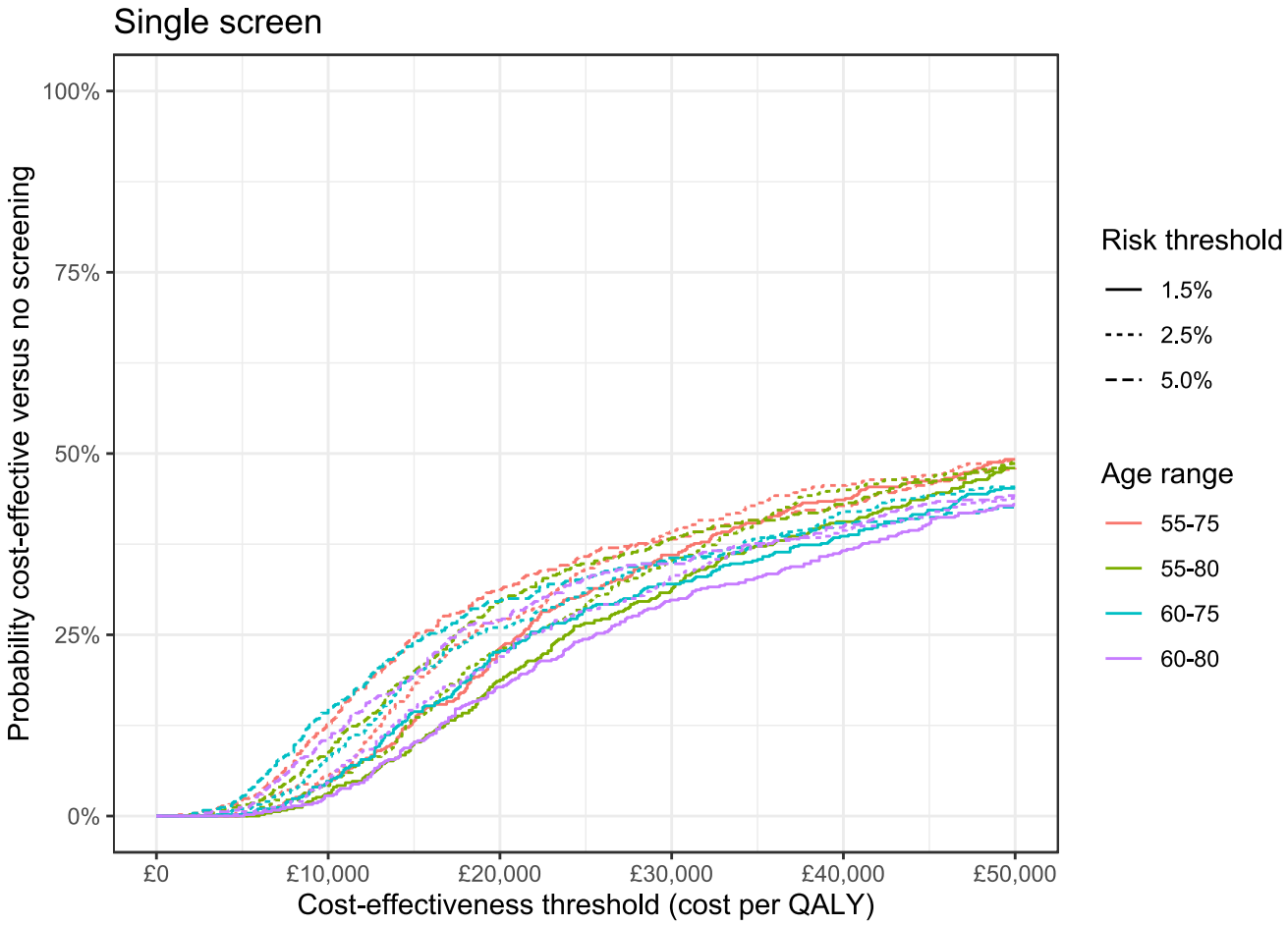


Figure 15 Probability of the single screening strategies being more cost-effective than no screening



Assuming an eligible population of 55-80 year olds at a predicted risk of lung cancer of ≥1.5%, Figure 16 shows the probability that any one screening schedule is the most cost-effective across a range of willingness to pay thresholds (£0 to £50,000 per QALY gained). For low WTP thresholds (<£10,000 per QALY gained), no screening is estimated as the most cost-effective strategy. However, as the WTP threshold increases, the annual screening strategy has >50% probability of being the most cost-effective strategy.

Figure 16 Cost-effectiveness acceptability curves for all strategies in a population aged 55-80 years old with ≥1.5% predicted risk of lung cancer

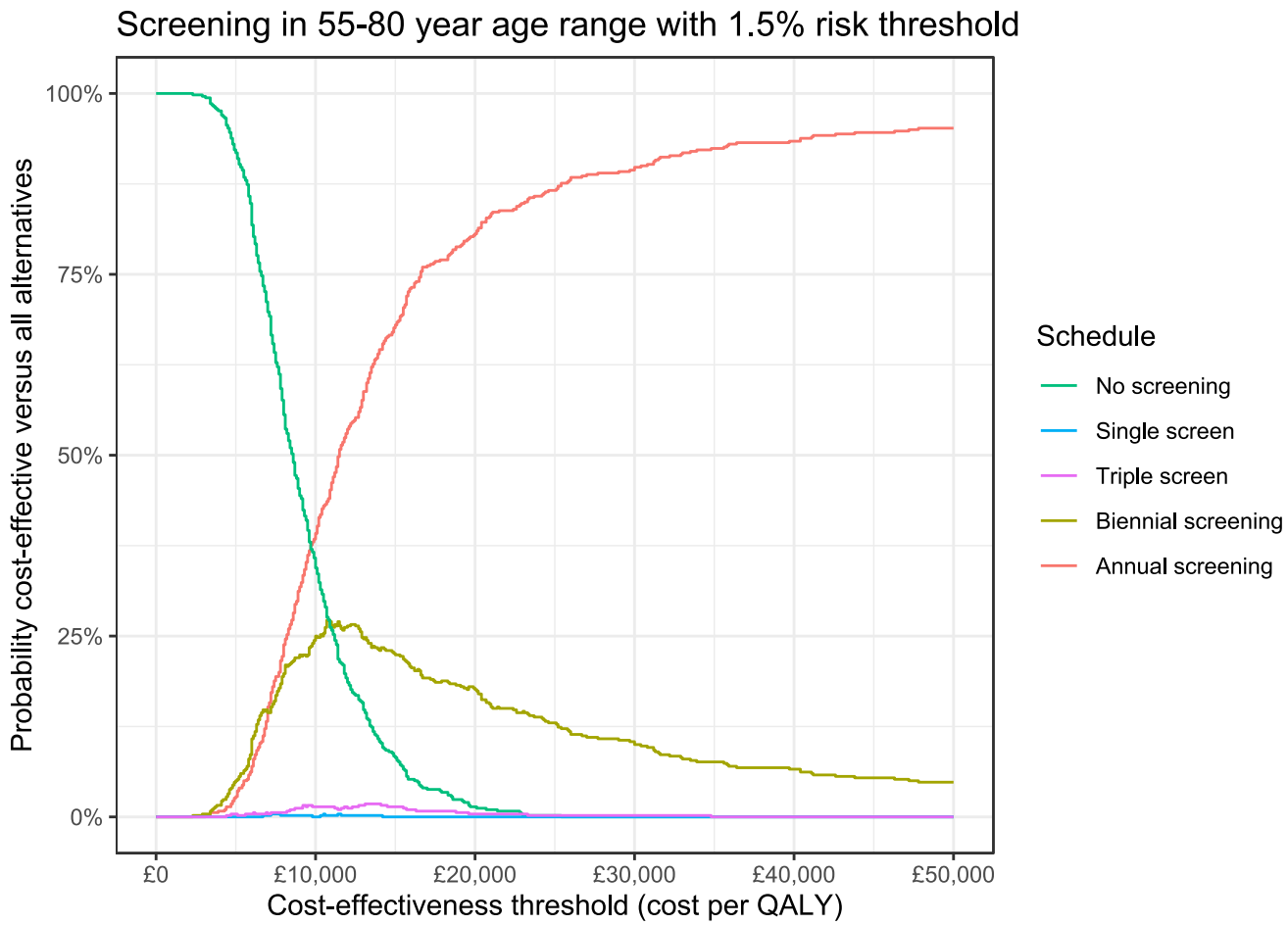


Table 23 Probabilistic base case results for strategies on the cost-effectiveness frontier (by decreasing INMB*)

Frontier strategy (ranked by decreasing INMB)	Lung cancer mortality in joiners	Number false positive screens per joiner	Life Years, undiscounted, ever-smokers 55-80	Lifetime QALYs, discounted, ever-smokers 55-80, [95% CrI]	Lifetime costs, discounted, ever-smokers 55-80, [95% CrI]	Lifetime INMB vs No screening at £20,000 per QALY gained, discounted, ever-smokers 55-80, [95% CrI]	ICER vs No Screening, [95% CrI]
No screening	-	-	17.349	9.793, [9.548 to 10.018]	£1,085, [£0,843 to £1,336]	-	-
A-55-75-1.5%	9%	0.43	17.378	9.805, [9.563 to 10.031]	£1,195, [£0,965 to £1,432]	£142, [£302 to £14]	£8,517, [£4,119 to £18,287]
A-55-80-1.5%	9%	0.39	17.379	9.806, [9.564 to 10.033]	£1,205, [£0,973 to £1,442]	£136, [£300 to £8]	£9,073, [£4,636 to £19,005]
A-55-75-2.5%	11%	0.41	17.370	9.802, [9.559 to 10.027]	£1,152, [£0,920 to £1,395]	£121, [£265 to £6]	£7,129, [£3,056 to £18,584]
A-55-75-5.0%	13%	0.37	17.361	9.798, [9.555 to 10.024]	£1,117, [£0,887 to £1,361]	£79, [£181 to -£8]	£5,837, [£1,392 to £24,571]

*calculated at a £20,000 per QALY gained willingness to pay threshold

In Table 23, only strategies on the cost-effectiveness frontier are shown, and are presented in terms of decreasing INMB. There were five strategies on the cost-effectiveness frontier (i.e., strategies that can give the maximum incremental net monetary benefit at a willingness to pay of £20,000 per QALY gained): no screening and four annual screening strategies, all with a lower age limit of 55 years.

Across the four screening strategies, the mean proportion of lung cancer deaths observed in individuals joining the screening programme ranged from 9% in the annual strategies with the lowest risk level, to 13% for the strategy targeted to those aged 55-75 years at the highest lung cancer risk level. The mean number of false positive LDCT screens experienced per person who joins the screening programme ranged from 0.37 for the strategy targeting individuals at the highest risk level (A-55-75-5%) and 0.43 for the A-55-75-1.5% strategy.

Across the whole population of ever-smokers aged 55-80 years (which includes individuals ineligible for screening, as well as those who do not take up screening), the mean per-person undiscounted life-years for the four screening strategies were estimated to be just over 17 years. Discounted mean lifetime QALYs across the whole population ranged from 9.798 to 9.806, with discounted mean lifetime costs between £1,117 and £1,205 per person. This compares to discounted mean QALYs and costs of 9.793and £1,085, respectively, for the no screening strategy. Strategy A-55-75-1.5% provides the highest INMB compared to no screening at a willingness to pay of £20,000 per QALY (£142), indicating it is the most cost-effective strategy.

In Figure 17, the incremental QALYs and costs for each of the 500 simulations in the probabilistic analysis for the four strategies on the cost-effectiveness frontier are shown. The triangles represent the mean across the simulations. Also presented are the results from a microsimulation analysis for each of these strategies. As can be seen, for the strategies targeted at individuals with lung cancers risks <5%, the deterministic results generally indicate fewer incremental QALYs gained and more incremental costs than results from the probabilistic analyses.

Figure 17 Scatter plot of the incremental QALYs and costs for the four annual screening strategies on the cost-effectiveness frontier for the probabilistic analysis

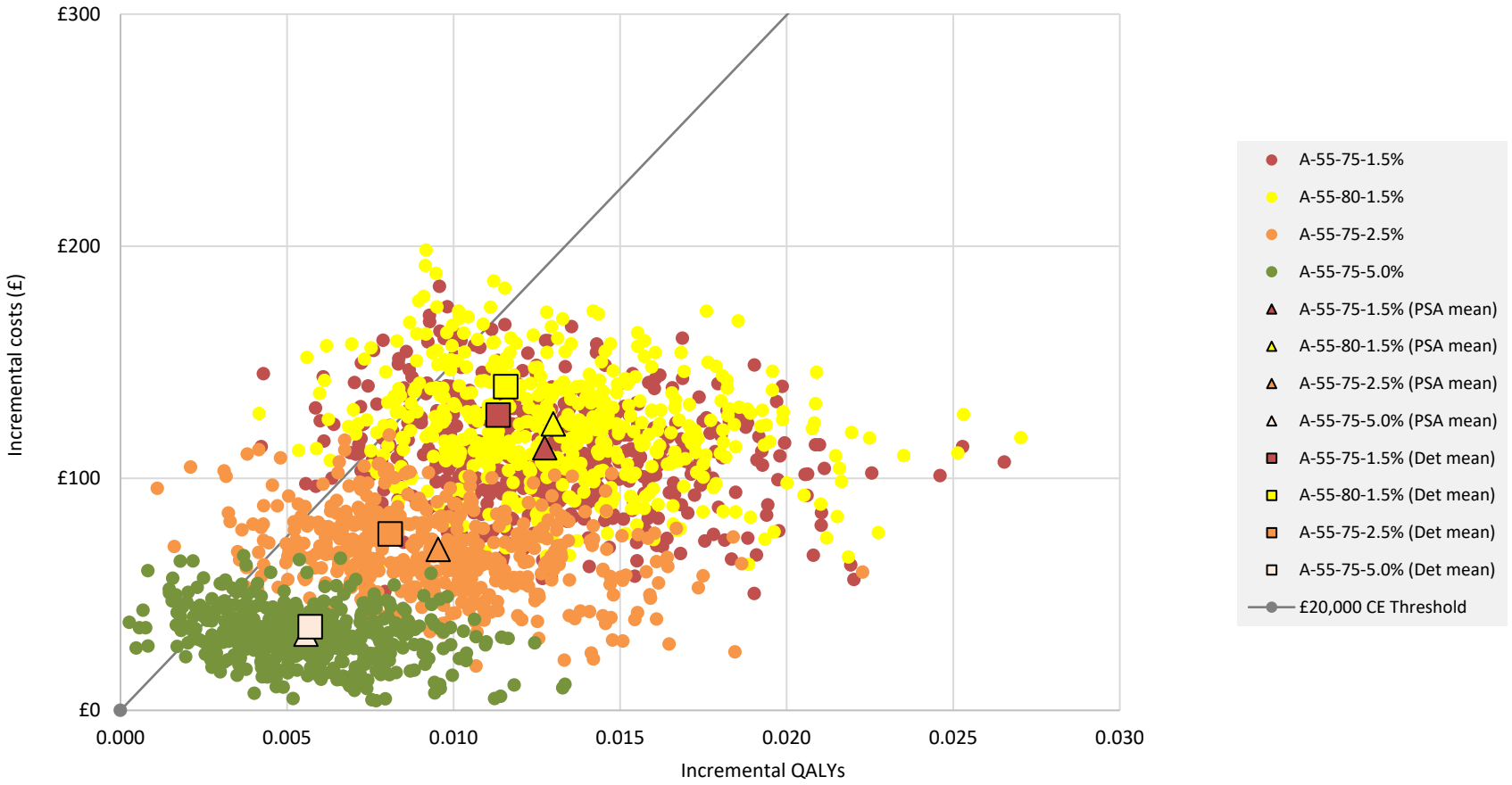
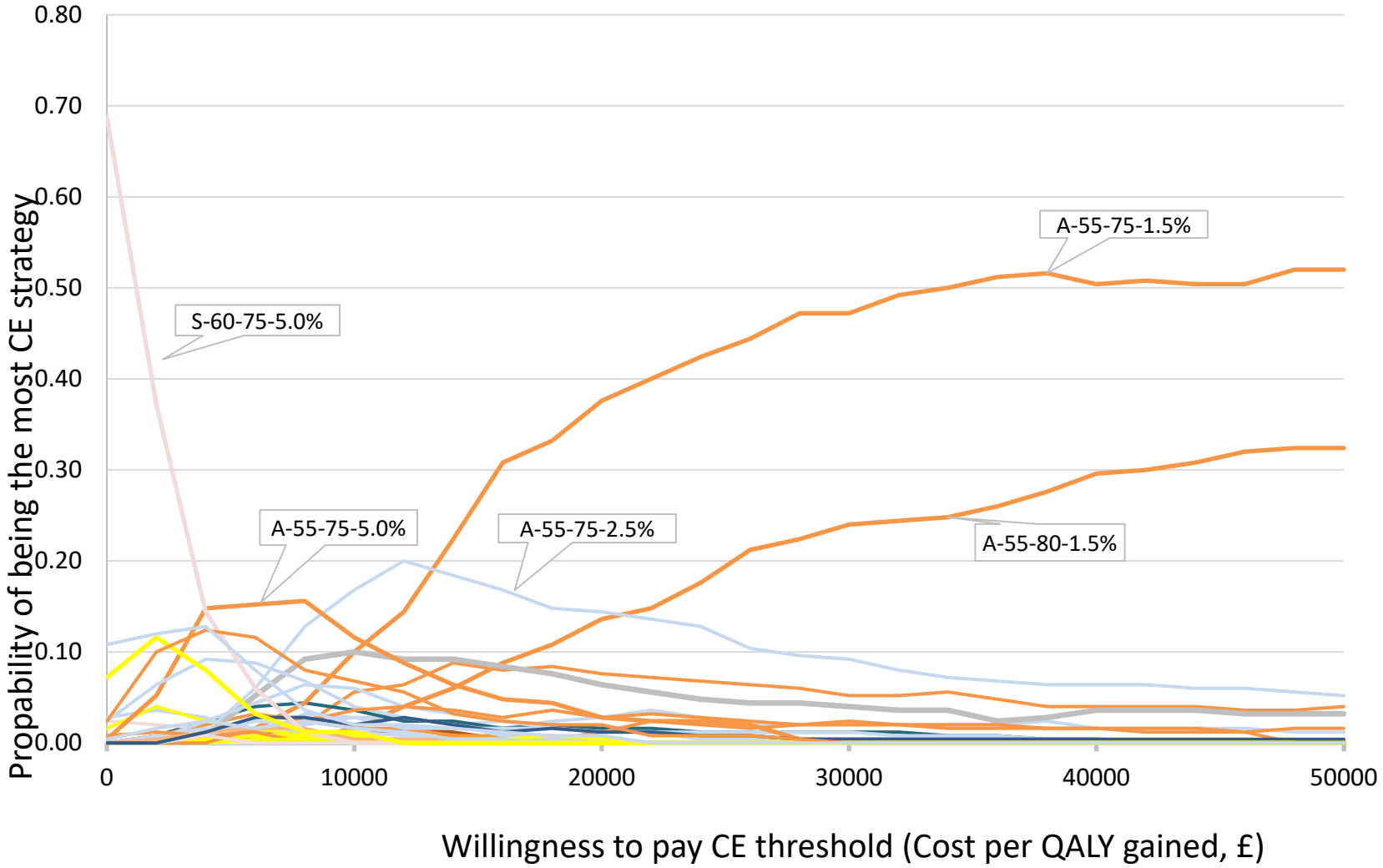


Figure 18 presents the cost-effectiveness acceptability curves for all screening strategies, showing the relative likelihood of each strategy being the single most cost-effective option across a range of willingness to pay thresholds (£0 to £50,000 per QALY gained). The figure highlights that at the £20,000 WTP threshold, the A-55-75-1.5% strategy had the highest probability (38%) of being the most cost-effective strategy. As the WTP threshold increased, the probability that this strategy was the most cost-effective also increased: above a WTP threshold of £30,000 per QALY, the A-55-75-1.5% strategy had a 50% probability of being the most cost-effective strategy. Two other strategies on the cost-effectiveness frontier each had a probability of approximately 15% of being the most cost-effective strategy at a WTP threshold of £20,000 per QALY. For very low WTP thresholds, say <£5,000 per QALY, a single screening strategy in a narrow age range at the highest risk threshold (S-60-75-5%) had the highest probability of being the most cost-effective strategy.

Figure 18 Cost-effectiveness acceptability curve



Microsimulation analysis

The results of a microsimulation analysis presented below were conducted by simulating a cohort of 3,000 individuals through the DES model and leaving all parameters at their central estimates.

The different population selection criteria produced a wide range of proportions of smokers joining screening programmes, with those strategies having the largest eligible age range and lowest risk thresholds having a higher proportion of joiners, as would be expected. Proportion of joiners across strategies ranged : from 2.5% for those aged 60–75 with ≥5% risk of lung cancer (60-75-5%) to 12.3% for those aged 55–80 with ≥1.5% risk of lung cancer (55-80-1.5%), as shown in Table 24. The predominant reasons for ever-smokers not joining screening programmes were not responding to the initial invitation or being of too low predicted risk of lung cancer.

Table 24 Percentage of population (ever-smokers 55-80 years) joining and not joining screening

	Proportion of base population aged 55-80 years (%)				
		Non-joiners			
Population criteria	Joiner	Decline	Risk too low	No response	Outside age/ Not invited
No screening	-	-	-	-	100.00
55-80-1.5%	12.3	2.4	36.1	49.2	0.0
55-80-2.5%	7.6	1.5	41.7	49.2	0.0
55-80-5.0%	3.4	0.7	46.7	49.2	0.0
60-80-1.5%	10.8	2.1	25.1	36.8	25.2
60-80-2.5%	6.7	1.3	30.0	36.8	25.2
60-80-5.0%	3.2	0.6	34.2	36.8	25.2
55-75-1.5%	10.5	2.1	34.6	45.8	7.0
55-75-2.5%	6.3	1.2	39.6	45.8	7.0
55-75-5.0%	2.7	0.5	43.9	45.8	7.0
60-75-1.5%	9.0	1.8	23.6	33.4	32.2
60-75-2.5%	5.4	1.1	27.9	33.4	32.2
60-75-5.0%	2.5	0.5	31.4	33.4	32.2

Table 25 presents the stage distribution of lung cancer diagnoses in those individuals who join the screening programme, regardless of whether these are diagnosed via clinical presentation or are screen-detected. For each frequency of screening, the average proportion of diagnoses across the different age and risk profiles is presented. Table 25 shows that screening strategies were associated with an increased probability of lung

cancer being diagnosed in the early stages (I and II) versus later stages (III and IV). Whereas 44% of NSCLC diagnoses are in stage IV in the no screening strategy, an average of 21% of NSCLC diagnoses are in stage IV when annual screening is implemented. Similarly, just 9% of NSCLC diagnoses are in stage IA for the no screening strategy, while 51% of NSCLC diagnoses are made in stage IA when annual screening is implemented.

Table 26 presents the stage distribution of screen-detected lung cancers only (thus there are no screen-detected cancers in the no screening strategy), in those individuals who join the screening programme. These estimates indicate that, depending on the screening frequency, at least 58% of screen-detected cancers are diagnosed in stage IA, with 73% of screen-detected cancers in annual screening strategies diagnosed in stage IA.

Table 25 Stage distributions of diagnoses by presentation or LDCT screening

	NSCLC							SCLC	
Screening design	IA1	IA2	IA3	IB	II	III	IV	Ltd	Ext
No screening	0%	5%	4%	5%	13%	28%	44%	30%	70%
Single	4%	8%	4%	5%	13%	27%	39%	30%	70%
Triple	7%	12%	4%	5%	12%	24%	37%	49%	51%
Annual	16%	27%	7%	3%	9%	16%	21%	65%	35%
Biennial	12%	20%	10%	4%	10%	18%	25%	59%	41%

Table 26 Stage distributions of diagnoses as detected by LDCT screening only

	NSCLC							SCLC	
Screening design	IA1	IA2	IA3	IB	II	III	IV	Ltd	Ext
No screening	0%	0%	0%	0%	0%	0%	0%	0%	0%
Single	28%	26%	5%	0%	10%	22%	10%	0%	0%
Triple	29%	32%	3%	0%	7%	15%	13%	100%	0%
Annual	24%	38%	10%	5%	6%	7%	9%	93%	7%
Biennial	21%	28%	17%	4%	7%	11%	10%	100%	0%

A comparison of benefits gained for the four screening strategies on the cost-effectiveness frontier is presented in Table 27. Individuals who join the a-55-75-5% screening strategy are estimated to receive the greatest gains in life-years and QALYs compared to no screening: 0.561 LYs and 0.209 QALYs. However, due to this strategy having only 2.7% of the population joining (see Table 24), the gains in life-years and QALYs across the whole population of ever smokers 55-80 years old was much diluted (0.015 LYs and 0.006 QALYs). On the other hand, strategy A-55-80-1.5% was associated with the fewest gains in those individuals joining (for strategies on the cost-effectiveness frontier), yet since 12.3%

of the population are estimated to join this strategy, the LYs and QALYs gained across the whole population are greater than that for the other strategies.

Table 27 Relative attainment of benefit between joiners and non-joiners for strategies on the cost-effectiveness frontier

Incremental benefit vs No Screening	Frontier strategy			
	A-55-75-1.5%	A-55-80-1.5%	A-55-75-2.5%	A-55-75-5%
Programme joiners				
Life-years gained	0.257	0.224	0.314	0.561
QALYs gained	0.108	0.094	0.128	0.209
Ever Smokers				
Life-years gained	0.027	0.027	0.020	0.015
QALYs gained	0.011	0.012	0.008	0.006

QALYs, quality-adjusted life year

Table 28 provides details on the clinical outcomes estimated by each of the four screening strategies on the cost-effectiveness frontier. The average number of screens over a life-time ranges is approximately 10. For the strategies with a 1.5% risk threshold, when the exit age is extended to 80 years of age, the average number of screens is lower than where the exit age is 75 (10.28 vs 11.39) due to those additional, older, individuals only meeting the eligibility criteria for a short period of time.

Across all strategies, between 58% and 67% of lung cancers are diagnosed via screening, with the highest estimate in the strategy having the highest risk threshold. Individuals receiving screening are estimated to have a mean of around 0.40 false positive LDCT screen results over the entire screening period. The estimation of the number of false positives is driven by the LDCT specificity estimate used in the model. Given that the mean LDCT specificity used in the model is 96.3% and individuals are estimated to have approximately 11 LDCT screens over their lifetime, the estimated number of false positive scans per person screened is entirely consistent.

Table 28 Clinical outcomes for programme joiners by strategy on the cost-effectiveness frontier

Secondary outcome	A-55-75-1.5%	A-55-80-1.5%	A-55-75-2.5%	A-55-75-5%
Mean number of screens per joiner	11.39	10.28	10.97	9.82
Proportion of diagnoses detected by screening (%)	60.0%	58.3%	65%	67%
Mean number of false positives* per joiner	0.42	0.38	0.42	0.38
Mean lead time (months)	12.50	11.87	13.72	15.91
5-year lung cancer survival (%)	19.0%	18.3%	20.2%	21.6%
Compared to No screening				
Change in lung cancer mortality (%)	-1.5%	-1.2%	-1.8%	-2.6%
Additional survival time with lung cancer (years)	2.99	2.84	3.29	3.61
Change in age at lung cancer diagnosis(years)	-2.51	-2.40	-2.82	-2.44
Change in age at death from lung cancer(years)	0.47	0.44	0.47	1.16
Change in age at death from other causes (years)	0.14	0.11	0.21	0.47
Per 100,000 programme joiners				
Number of screen-detected cases	25,652	23,860	32,226	46,391
Number of interval cancers	1,583	1,356	1,585	2,442
Additional lung cancer diagnoses (compared to no screening)	7,284	7,050	10,037	13,429
Lung cancer deaths averted	3,484	2,711	4,226	6,104

*false positives defined as those individuals referred to MDT but do not subsequently receive a diagnosis of lung cancer

The lead time is the difference in time between diagnosis of lung cancer by screening and diagnosis of that same lung cancer (or death from other causes, whichever comes first) if the individual had not participated in screening. Thus, for the A-55-75-5% strategy simulated individuals diagnosed with lung cancer were detected at a mean of 15.9 months earlier than they would have been detected (or died from other causes) in the no screening strategy. Strategy A-55-80-1.5% is associated with the lowest mean lead time, but includes older individuals (who are more likely to die from other causes). The trend for mean lead time to increase as the risk threshold for the strategy increases highlights the greater benefit seen in identifying those at greater risk. However, as Table 27 shows, once the benefits are averaged across the whole population invited for LDCT screening, the benefits are diluted.

Screening strategies on the cost-effectiveness frontier are associated with five-year lung cancer survival estimates between 18% to 22%, reductions in lung cancer mortality of 1.2% to 2.6% (in those joining the screening programme), and between 2.84 and 3.61 additional years of life with lung cancer compared to no screening. Overall, strategy A-55-80-5% is associated with better lung cancer survival outcomes in those who join the programme compared to the other strategies on the cost-effectiveness frontier. This is due to the higher

risk threshold, meaning that it is more likely to identify more lung cancers (as results in the bottom section of Table 28 indicate), therefore have more opportunity for benefit.

Table 29 shows the total number of lung cancers diagnosed within each of the screening strategies on the cost-effectiveness frontier compared to no screening in their specific age and risk-defined subgroups, per 100,000 individuals joining the programme.
The number of lung cancer diagnosed shown when no screening programme is implemented are different across the strategies because we are only interested in the subgroups that meet the different eligibility criteria for the strategies. For example, for comparison with strategy A-55-75-1.5%, we are only considering lung cancer diagnoses in those individuals aged 55-75 years and have a risk of lung cancer of ≥1.5% in the absence of screening. The number of lung cancers observed is highest in the A-55-75-5% group, with screening leading to an additional 13,429 lung cancer diagnoses, per 100,000 individuals joining the programme, a relative increase in diagnoses of 25.3%. The higher number of diagnoses observed with screening in the higher risk strategies, reflects that these individuals are at risk of dying from other causes before presenting with lung cancer.

Table 29 Lung cancer diagnoses (NSCLC and SCLC) in screening strategies on the cost-effectiveness frontier, per 100,000 joining participants

Strategy	Lung cancers diagnosed when a screening programme is implemented	Lung cancers diagnosed when no screening programme is implemented	Additional lung cancer diagnoses from implementation of screening programme	Relative risk of lung cancer diagnosis
A-55-75-1.5%	42,753	35,469	7,284	1.205
A-55-80-1.5%	40,941	33,892	7,050	1.208
A-55-75-2.5%	49,659	39,622	10,037	1.253
A-55-75-5.0%	69,586	56,157	13,429	1.239

In Table 30, the number of lung cancer deaths per 100,000 individuals who join the screening programme are presented. As in Table 30 above, the number of lung cancer deaths shown when no screening programme is implemented are different across the strategies because we are only interested in the subgroups that meet the different eligibility criteria for the strategies. Strategy A-55-75-5% is associated with the largest number of deaths averted compared to no screening, a reduction in lung cancer mortality of 13.9%. These findings are consistent with Table 27 and Table 28, where LDCT screening in a population at high risk leads to more benefits for that population than if a lower risk threshold is assumed. However, as pointed out, once all benefits are averaged across the total eligible population, the benefits are much diluted.

Table 30 Lung cancer deaths (NSCLC and SCLC) in strategies on the cost-effectiveness frontier, per 100,000 joining participants

Strategy	Lung cancer deaths with screening programme	Lung cancer deaths without screening programme	Averted lung cancer deaths	Relative risk of lung cancer death
A-55-75-1.5%	24,385	27,869	3,484	0.875
A-55-80-1.5%	23,589	26,300	2,711	0.897
A-55-75-2.5%	26,414	30,641	4,226	0.862
A-55-75-5.0%	37,845	43,949	6,104	0.861

The costs associated with the screening strategies on the cost-effectiveness frontier are shown in Table 31. The costs per joiner that specifically relate to the LDCT scans for individuals joining the screening programmes on the cost-effectiveness frontier ranged from £883 (A-55-75-1.5%) to £776 (A-55-80-5%). The lower costs associated with LDCT scans in A-55-75-5% reflect the lower number of LDCT scans in this population. This is because there are more lung cancers identified and individuals with a diagnosis will not continue to attend screening.

Lung cancer costs consisting of diagnosis, treatment and follow-up costs, were highest for the strategy, with highest risk threshold (A-55-75-5%), the largest difference seen in the diagnostic costs. Higher lung cancer related costs in the A-55-75-5% strategy are explained by the greater chance of identifying a lung cancer in this population, because individuals are at a higher risk. The screening programmes on the cost-effectiveness frontier are estimated to lead to population lifetime cost increases of £469M to £1,814M for a relevant population of 13 million ever-smokers aged 55–80 years.

Table 31 Costs associated with screening strategies on the cost-effectiveness frontier

	Frontier strategy			
Cost item	A-55-75-1.5%	A-55-80-1.5%	A-55-75-2.5%	A-55-75-5%
Per Ever-smoker aged 55-80 years				
Screening programme admin	£4.70	£5.06	£4.66	£4.63
Total lifetime cost (incl. admin)	£1,298	£1,311	£1,247	£1,207
Per Programme joiner (lifetime)				
LDCT screens	£883	£804	£854	£776
Lung cancer intervention (diagnosis)	£1,164	£1,161	£1,406	£2,123
Lung cancer intervention (recurrence)	£889	£906	£888	£1,372
Lung cancer intervention (follow-up)	£124	£114	£156	£206
End of life	£326	£323	£365	£559
TOTAL	£3,387	£3,307	£3,668	£5,036
Ever-smoker population aged 55-80 (Est. 13m)				
Screening administration (£,m)	£61	£66	£61	£60
LDCT screens (£,m)	£1,208	£1,285	£700	£275
TOTAL (£,m)	£16,878	£17,037	£16,212	£15,692
Additional cost vs No screening (£,m)	£1,654	£1,814	£988	£469

Controlled sensitivity analyses

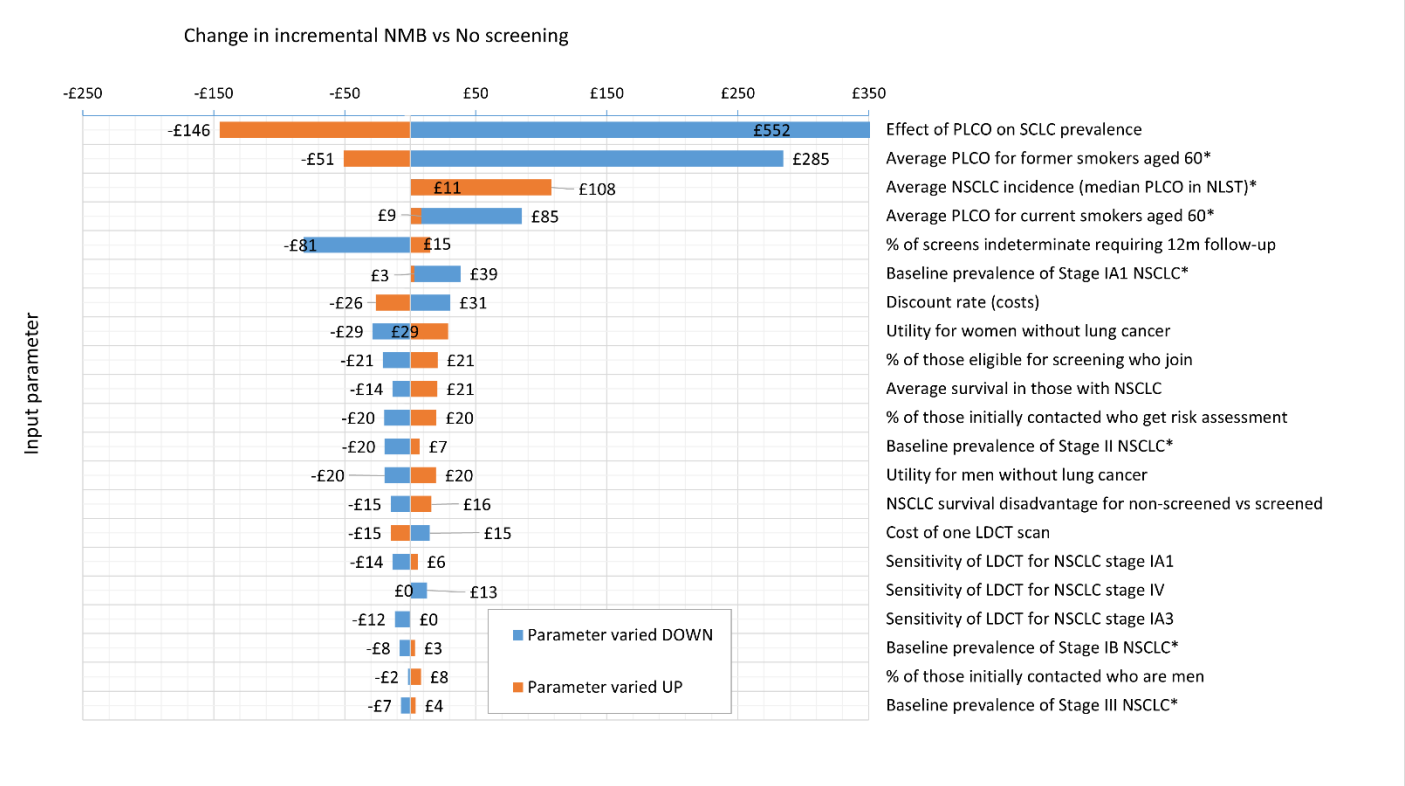
In this section, controlled one-way sensitivity analyses and scenario analyses are presented. All of these analyses are reported for strategy A-55-75-1.5% which was the optimal screening strategy in probabilistic base case analyses. The results are reported in terms of the INMB associated with A-55-75-1.5% compared to no screening at a willingness to pay threshold of £20,000 per QALY gained. The results for all of these sensitivity analyses are based on 3000 simulated individuals, as in the microsimulation analysis. We also used the same random number stream for each sensitivity analysis (which determines the latent times of events, and whether a screening test accurately classifies a screening participant), to reduce stochastic variation.

One-way sensitivity analyses

In the one-way sensitivity analyses, 40 parameters were individually increased and decreased, in turn, by 20%. This is an arbitrary value, which is not linked to the precision of the individual parameters, and does not maintain any correlations between parameter values. However, it does allow some assessment of the likely impact of differing values on results. Figure 19 is a tornado diagram showing the change in INMB for the 21 parameters

having the largest impact when increased/decreased by 20% on the INMB. In the probabilistic base case analysis the expected INMB for the strategy A-55-75-1.5% was £142, so a loss of INMB of this magnitude would render screening no longer cost-effective.

Figure 19 Tornado diagram for results of one-way sensitivity analysis for strategy A-55-75-1.5%



Parameters sorted by descending incremental NMB impact range; top 21 most impactful presented only.

There is some asymmetry in the tornado diagram (Figure 19) with the microsimulation result (represented by change in INMB of £0) not falling within the bounds of some of the resulting ranges (e.g. “Average PLCO incidence”, “Average PLCO for current smokers aged 60”); this indicates some non-linearity in the model.

The results indicate little sensitivity in the parameters to an arbitrary 20% change, with only a change in one parameter leading to an INMB that would not be considered cost-effective at the £20,000 willingness to pay threshold. This is the first parameter in Figure 12: “Effect of PLCO on SCLC prevalence”. This parameter determines how predictive baseline PLCOm2012 is for the prevalence of SCLC at baseline (which also indirectly affects the prevalence of NSCLC at baseline). As shown in Table 32 when this parameter is lowered

by 20% to 0.744, it substantially reduces the risk of individuals having SCLC at baseline, particularly in a targeted group. In contrast, when this parameter is increased by 20% to 1.116, the risk of SCLC at baseline increases substantially. This is important because participants with SCLC at baseline are less likely to benefit from screening than participants without SCLC at baseline.

Table 32 Impact of varying the parameter "Natural history prevalent state at entry, PLCO coefficient SCLC"

Prevalence at baseline	PLCO coefficient for SCLC prevalence		
(per 100,000)	Central estimate (0.930)	Lower estimate (0.744)	Upper estimate (1.116)
Full population			
NSCLC	629.5	630.2	628.3
SCLC	51.1	46.3	61.1
Aged 55-75, PLCO predicted risk ≥1.5%			
NSCLC	1571.3	1573.3	1567.8
SCLC	130.4	106.3	167.2

The 95% credible interval for this parameter is (0.709, 1.204). Note, however, that this parameter is negatively correlated with two other parameters which also determine the prevalence of SCLC at baseline. If these parameters take their expected value conditional on the PLCO coefficient taking the lower and upper estimates, the range of SCLC prevalence at baseline contracts to 49.1–57.5 per 100,000 for the full population and 112.7–157.2 per 100,000 for the targeted population.

Scenario analyses

Twenty scenario analyses were conducted, in which changes to the model structure, or an individual or set of parameter values were made. As above, these analyses were based on strategy A-55-75-1.5%. The results were also reported for the strategy giving the highest INMB versus no screening of all screening strategies, if not strategy A-55-75-1,5%. The results of these scenario analyses are presented in Figure 20 and Table 33.

Figure 20 Change in INMB from the base case analysis for the scenario analyses for strategy A-55-75-1.5%

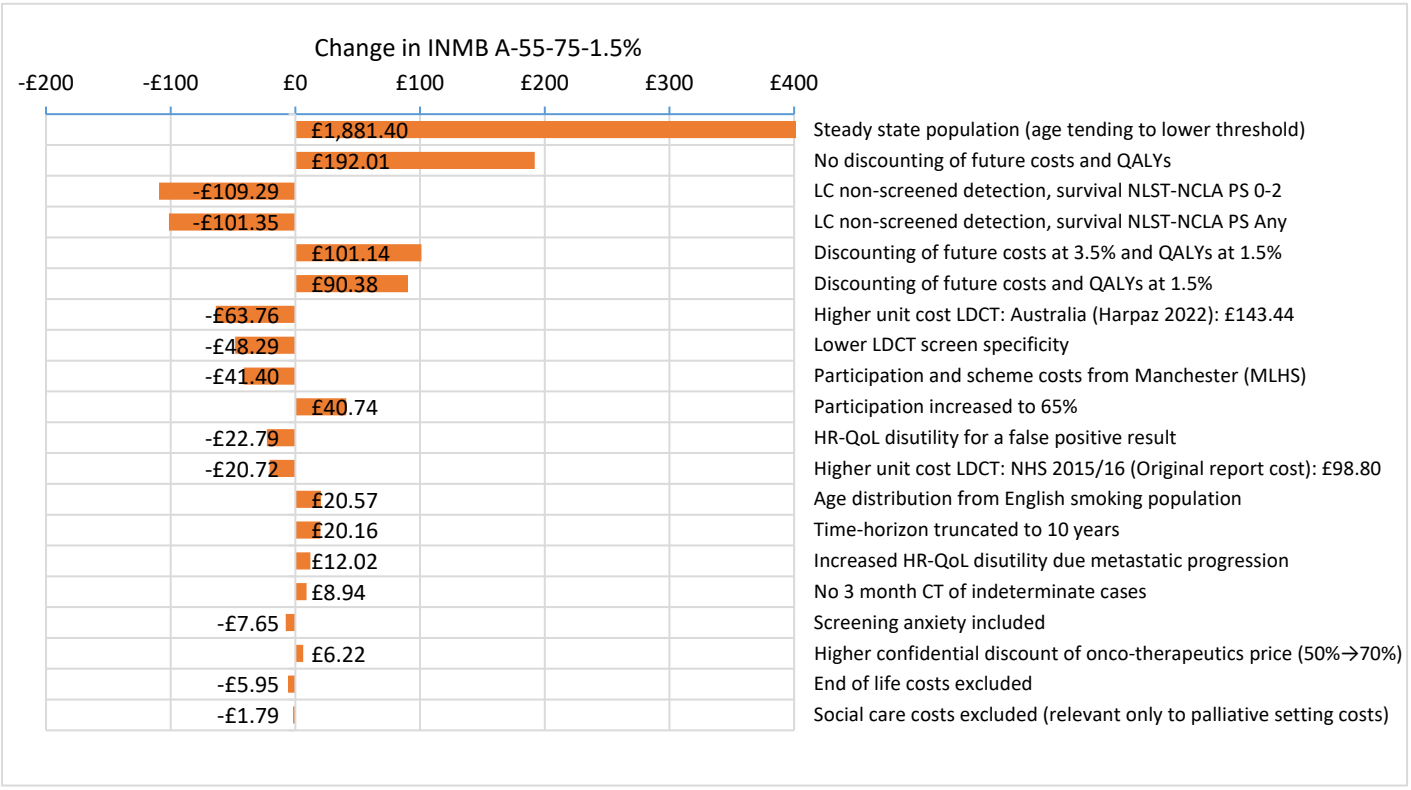


Table 33 Change in INMB from the base case analysis for the scenario analyses for strategy A-55-75-1.5%

Scenario	Change in INMB* vs No Screening from base case analysis	Strategy in scenario with the highest INMB*	
Steady state population (age tending to lower threshold)	£1,881	A-55-80-1%	£2,077
No discounting of future costs and QALYs	£192	A-55-75-1.5%	£192
LC non-screened detection, survival NLST-NLCA PS 0-2	-£109	S-55-75-1.5%	-£16
LC non-screened detection, survival NLST-NLCA PS Any	-£101	S-55-75-1.5%	-£16
Discounting of future costs at 3.5% and QALYs at 1.5%	£101	A-55-75-1.5%	£101
Discounting of future costs and QALYs at 1.5%	£90	A-55-75-1.5%	£90
Higher unit cost LDCT: Australia(68): £143.44	-£64	B-55-80-5.0%	-£16
Lower LDCT screen specificity	-£48	B-55-80-5.0%	-£14
Participation and scheme costs from Manchester (MLHC)	-£41	A-55-75-1.5%	-£41
Participation increased from 51% to 65%	£41	A-55-75-1.5%	£41
HR-QoL disutility for a false positive result	-£23	B-55-80-5.0%	-£11
Age distribution from English smoking population	£21	A-55-75-1.5%	£21
Higher unit cost LDCT: NHS 2015/16 (Original report cost): £98.80	-£21	B-55-80-5.0%	-£10
Time-horizon truncated to 10 years	£20	A-55-75-1.5%	£20
Increased HR-QoL disutility due metastatic progression	£12	A-55-75-1.5%	£12
No 3 month LDCT of indeterminate cases	£9	A-55-75-1.5%	£9
Screening anxiety included	-£8	A-55-75-1.5%	-£8
Higher confidential discount of onco-therapeutics price (50%→70%)	£6	A-55-75-1.5%	£6
End of life costs excluded	-£6	A-55-75-1.5%	-£6
Social care costs excluded (relevant only to palliative setting costs)	-£2	A-55-75-1.5%	-£2

*at a willingness to pay threshold of £20,000 per QALY gained. HR-QoL, health-related quality of life; INMB, incremental net monetary benefit; LDCT, low dose computed tomography; MLHC, Manchester Lung Health Check; NLCA, national lung cancer audit; NLST, National Lung Screening Trial; PS, performance status; QALYs, quality-adjusted life years; YLST, Yorkshire Lung Screening Trial

In summary, the scenario analysis results support the robustness of the outcome indicated by the one-way sensitivity analysis. Incremental net monetary benefit (at 3,000 simulations) remains positive in all but two cases for A-55-75-1.5%, at a WTP threshold of £20,000 per QALY gained. These two cases relate to calculation of non-screen detected lung cancer (see below for further details on these scenarios).

The strategy delivering the highest net monetary gain over no screening (A-55-75-1.5%) is unchanged from the base case in all but seven of the twenty scenarios when a WTP threshold of £20,000 per QALY gained is assumed. When lower LDCT screen specificity is assumed, a HR-QoL disutility for a false positive result is assumed, of higher LDCT costs are assumed, strategy B-55-80-5% is estimated in these deterministic analyses are the optimal screening strategy. On the other hand, when a smaller impact on lung cancer survival from screening is assumed, a single screen strategy (S-55-75-1.5%) is estimated as the most cost-effective strategy. When a steady state population is assumed, an annual strategy for the largest group of eligible individuals (A-55-75-1.5%) is the optimal strategy with an increase in INMB of £2,077 (see below).

Steady state population

This scenario explores the impact of entry age distribution skewing heavily left towards minimum entry age in the range, as the screening eligible population mean age reduces as years pass since inception of the programme Only the 55-80 year and 60-80 year ranges were tested, with all individuals assumed to be aged 56 and 61 years respectively. This approximates a steady-state model. The initial result found that too few individuals met the minimum risk thresholds at 3%, 4% and 5% given the younger age range at entry, therefore risk thresholds were reduced to 1%, 2% and 3%. In this scenario, A-55-80-**1%** had the greatest change in incremental net monetary gain (£2,077), indicating that once most eligible individuals are at the entry age (after a screening programme has been running for a number of years), annual screening at a low risk threshold until age 80 years would be cost-effective.

Discount rates

Adjustment of discount rates for future costs and benefits leads to increases in the INMB of strategy A-55-75-1.5%, and it remains the most cost-effective screening strategy. The removal of the annual discounting of future costs and benefits (3.5% in the base case) increases the INMB for strategy A-55-75-1.5% by £192. Reducing the discount rates for both to 1.5% increases the INMB by £90, while differential discounting rates (benefits at 1.5% and costs at 3.5%) leads to an increase of £101 in INMB.

Survival for non-screen detected lung cancers

In the base case analysis NLST data were used to inform survival associated with non-screen detected lung cancers. As noted in the Methods section, use of NLST data may lead to overestimates of the impact of LDCT screening on survival. This is because it was observed that survival for very early stage screen-detected cancers from NLST was greater than that observed in limited UK data, while survival for non-screen-detected stage IV cancers were estimated to be lower from NLST than observed in NLCA data. Thus, use of NLST may lead to a larger difference in survival between screen-detected and non-screen-detected cancers than would be observed in the UK. Thus, limited UK data were used in scenario analyses to explore this potential impact. Two scenarios were run using the UK data, one where it is assumed all individuals have PS 0-2, the other for individuals across the range of PS. The analyses indicate that in these two scenarios strategy A-55-75-1.5% would not be considered cost-effective at a willingness to pay threshold of £20,000 per QALY gained. Instead, a single screen strategy would be considered the most cost-effective (S-55-75-1.5%). However, if a willingness to pay threshold of £30,000 per QALY gained – the upper limit used by NICE – then strategy A-55-75-1.5% would be considered cost-effective, with INMB of £38 (for survival based on a population with PS 0-2) and £50 (for survival based on the whole population). Thus, as would be expected, if a more pessimistic assumption of the impact of screening on survival is assumed (as in the scenario analyses), screening is not considered as cost-effective.

Computed tomography screening costs

Two scenario analyses explored the impact of increasing the cost for LDCT scans. One was based on the cost used in the original model(5) (£98.80), the other taken from the recent evaluation of LDCT screening in Australia(68) (£143.44). As would be expected, use of higher cost estimates results in reduced INMB associated with strategy A-55-75-1.5%. Moreover, a biennial screening strategy with a higher risk threshold was estimated to be the most cost-effective in both scenarios: B-55-80-5.0%.

LDCT specificity

When the specificity of LDCT was reduced to match that reported by NELSON, the INMB for A-55-75-1.5% reduced by £48, and the biennial strategy B-55-80-5% was estimated to be the most cost-effective option with a reduction in INMB of £48.

Programme uptake

When programme uptake was matched to the Manchester LHC recruitment programme, i.e., lower rate of response, higher uptake, and face to face risk assessment (higher cost),

the INMB for A-55-75-1.5% reduced by £41. When a higher participation rate (65% rather than 51%) is tested, the INMB for S-55-75-1.5% increased by £41. In both scenarios A-55-75-1.5% remained the most cost-effective option.

Impact on health-related quality of life

Three scenario analyses focussed on health-related quality of life, that there would be an impact on HRQoL for a period following a false positive result, an increase in the disutility for stage IV lung cancers and a disutility experienced in the run-up to every screen representing screening anxiety. Scenarios where disutility for screening or a false positive result lead to a reduction in the INMB for strategy A-55-75-1.5% by £23 for false positives and £8 for screening anxiety. When disutility for false positives was assumed, a biennial strategy (B-55-80-5.0%) was estimated as the most cost-effective with a reduction in INMB of £13. When a larger disutility for stage IV lung cancer was assumed the INMB for A-55-75-1.5% increased by £12, and remained the most cost-effective strategy.

Age distribution

This scenario involved a change from the UKLS age distribution of responders to the age distribution of smokers in the UK population. This change led to an increased in the INMB for strategy A-55-75-1.5%, of £20, and it remained the most cost-effective strategy.

Time horizon

Truncating the time-horizon to 10 years increased the INMB for A-55-75-1.5% by £20.

Indeterminate cases

In this scenario, we excluded the 3 month cautionary follow-up LDCTs performed in a proportion of individuals who have an indeterminate LDCT scan result (LDCTs at 12 months follow-up are retained even if not already scheduled via the programme). The overall impact is very slight increasing the INMB for A-55-75-1.5% by £9.

Commercial discounting of higher cost drugs

In this scenario of the estimated confidential commercial discount of high cost lung cancer drugs was increased from 50% to 70%. The INMB for strategy A-55-75-1.5% increases slightly by £6, and the strategy remains the most cost-effective.

Social care and End-of-life costs

When all end-of-life costs for lung cancer deaths were excluded, the INMB for A-55-75-1.% reduces by £6. When costs relating to social care were excluded (approximately one-third of lung cancer end-of-life costs) there was very little impact on the INMB for A-55-75-1.5%

(decrease of £1). In both scenarios A-55-75-1.5% remained the most cost-effective strategy.

Discussion

Main findings

- All screening strategies lead to increased costs and QALYs gained compared to no screening strategy
- All targeted screening strategies except those using only a single screen have positive INMB at £20k per QALY. This is equivalent to saying that their cost per QALY is below a threshold of £20,000
- For instance the INMB for an annual strategy inviting 55-75 year olds and screening those with a risk greater than 1.5% (A-55-75-1.5%) was £142, [95% Credible Interval £302 to £14]. The cost per QALY was £8,517, [95% Credible Interval £4,119 to £18,287]
- Triple screen strategies are marginally cost-effective compared to no screening, producing small INMB an order of magnitude lower than those available via targeted annual screening
- Annual strategies are generally associated with the greatest incremental costs (i.e., greatest budget impact and use of LDCT) and QALYs gained, with single screening strategies leading to lowest incremental costs, but also lowest incremental QALYs.
- Although biennial screening strategies are cost-effective versus no screening, they are not cost-effective in a fully incremental cost-effectiveness analysis – this assumes, though, that the opportunity cost per LDCT scan is the same for an annual screening strategy as a biennial strategy, which may not be a valid assumption if LDCT scanners and/or radiologist time are highly constrained resources
- Strategies on the cost-effectiveness frontier from the prob analyses are all annual strategies, with the strategy having the widest inclusion criteria (age 55-75 years at 1.5% risk of lung cancer) estimated to lead to the greatest INMB gained.
- There is a trade-off in the benefits gained for those joining the screening strategies, and the % of the population that are eligible for screening. Thus we are likely to see the greatest benefits for those joining programmes when we restrict screening to high risk individuals, but when gains are spread across the total eligible population, such strategies are not as attractive and screening strategies targeting lower risk populations preferred

Strengths

- The model was developed by an independent research group
- There was considerable external input from clinical and methodological experts
- The model has evolved through three major iterations over 6 years each iteration being exposed to peer review
- It is based on the most up-to-date available evidence applicable to the NHS
- It includes a completely new natural history model that better reflects the expected stage distributions at diagnosis with LDCT screening

    - We have conceptualised, built and calibrated a natural history model for lung cancer which is suitable for evaluating lung cancer screening programmes. It takes into account that different individuals have different risks for lung cancer (depending on their smoking history and other risk factors), and that different cancers progress at different rates. It is therefore ideally constructed to support evaluations of targeted lung cancer screening programmes (where the population is enriched according to risk factors), and when there is concern about overdiagnosis (extremely slow-growing cancers which would never be clinically meaningful in the absence of screening).
- It incorporates a large number of one-way sensitivity analyses and scenario analyses. The model behaves predictably in these and the results are robust to the changes examined in the scenario analyses
- The model can be updated with new evidence and used to address additional questions in the future on the implementation of LC screening

Limitations

- Natural history model:
    - There are certain mathematical assumptions which make the natural history more computationally tractable, e.g., all transitions between health states representing progression and presentation are assumed to have a constant hazard rate, although the assumption of heterogeneity in progression rates counteracts this to some degree. It is likely not feasible to empirically test whether these assumptions are realistic, although other models (e.g., a model of tumour volume doubling) could suggest alternative assumptions.
    - Due to computational issues with the Hamiltonian Monte Carlo sampler, we used a surrogate for the likelihood function which was more computationally tractable. This may have introduced some bias (though there is no clear evidence of the model being unfit for purpose) and future work will aim to address this.
    - Only includes data from NLST – technology may have improved, technology may be used differently now and in the UK, healthy volunteer effect, 8th edition stages needed to be imputed
    - Would ideally also incorporate evidence from other studies, e.g., NELSON and UK-based lung cancer screening pilots

- Results:

    - Demonstrate a great deal of Monte Carlo variation when 3000 simulants are modelled (deterministic vs probabilistic results) – this needs to be considered when examining the scatter plots

- Cost of LDCT may not reflect true opportunity cost. Issues to consider include:
    - Would a screening programme displace other patients from using LDCT capacity (leading to delays)?
    - If radiologist time for interpretation comes from current contracts what does this actually displace?
    - If radiologists are paid to do interpretation outside their current contracts, does this incur a higher rate, and does it in fact still displace current NHS activity?

- There is limited evidence for
    - Survival from LC in UK
    - Other cause mortality and how it relates to smoking history
    - Utility associated with screening and having a FP scan, but also having
    - Estimation of PLCOm2012 in UK-based population, requiring number of assumptions

- Doesn’t account for
    - Increased cancers from radiation exposure
    - Costs or benefits associated with the reporting of incidental findings from the LDCT screens
    - Different ways incorporating smoking cessation into the screening programmes

Summary

Conclusions and implications for policy

After three major iterations and wide consultation over 6 years we have a developed a cost-effectiveness model for targeted lung cancer screening using LDCT (ENaBL), which has been validated and works well. It is robust in sensitivity analyses. In combination we thus have confidence in its results.

The results confirm earlier estimates from ENaBL that targeted LDCT LCS is cost-effective at £20,000/QALY, the lower threshold used by NICE. The INMB estimates valuing a QALY at £20,000 are positive consistent with this.

Our analysis suggests that the most cost-effective use of NHS resources involves a broad target population (adults aged 55–74 with a predicted risk of lung cancer in the next 6 years of ≥1.5%), which is very much in line with the target populations of UK-based pilots of LDCT LCS. Further, the analysis suggests that it is necessary to conduct *multiple* screens in the targeted population: a single screen will not be cost-effective, three screens will be marginally cost-effective, while prolonged annual screening will be the most cost-effective (although some potential harms including radiation-induced cancers have not been included in the analysis). Prolonged screening every two years (biennial LCS) can also lead to significant benefits and may be an appropriate alternative to annual screening if resources are too constrained. Alternatively, the risk threshold can be raised to reduce the population targeted for screening. We have not investigated the possibility of stratified screening, e.g., biennial LCS for those with a moderate risk of lung cancer and annual LCS for those with a high risk.

This report is already long and complex, but the model can produce further analyses to help policy makers understand the nature of the additional costs and benefits involved in more intensive approaches and relate the findings of the model to the findings of clinical studies. There may also be other analyses which we can present to fully explore the capability of the model to answer current policy questions concerning targeted LDCT LCS. Provided these were prioritised and limited in number these could be delivered as part of the current project.

For the future, with further investment and development the ENaBL model has the capacity to:

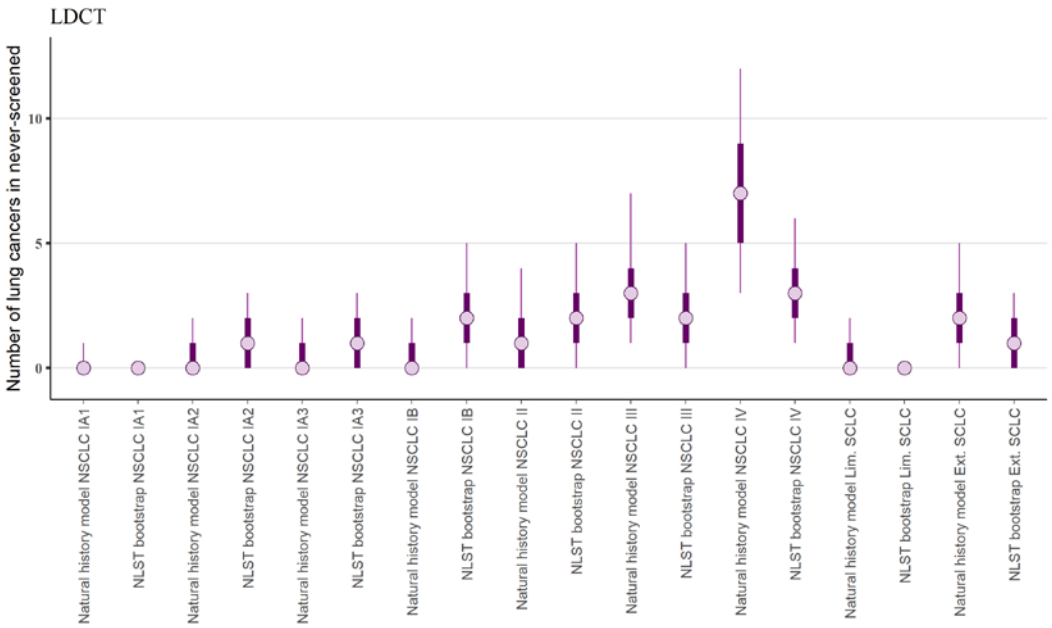
- Help with the assessment of feasibility and budget impact of introducing targeted LDCT LCS into health services, particularly the NHS. As well as modifying the model this would also require collation of data from piloting.
- Inform the cost-effectiveness of currently discussed adjuncts to targeted LDCT LCS such as incorporation of smoking cessation and more intensive response to incidental findings. There is already a project which intends to look at the cost-effectiveness of alternative approaches to risk assessment which has asked to use the final version of the ENaBL model when available. Stratification of screening so that different follow-up is offered depending on the results of the initial scan is another aspect of approach which is currently under discussion.
- Inform the effect on effectiveness and cost-effectiveness of future possible modifications to a targeted LDCT LCS. Modifications which are frequent in the evolution of national screening programmes include changes in screening interval and target group or the emergence of new screening tests. We expect that such changes will also affect targeted LDCT LCS if it is implemented. The ENaBL model could be the starting point for investigating impact of any modifications. As potentially the first targeted screening programme, relaxing the targeting of LDCT LCS programme could be a particularly pertinent and challenging future modification.
- Inform the effect on effectiveness and cost-effectiveness of other future changes. What for instance might be the effect of continuing reductions in rates of smoking, and so the risk of lung cancer?
- If implemented, monitor the effect on cost-effectiveness once the screening programme is in place. What for instance might be the effect of coverage falling or if there was difficulty engaging with all sections of the population at risk? This would be an extension of existing quality assurance processes.

Due to this update being part funded by the National Institute of Health Research (NIHR), we are prohibited from making policy implications. This is the role of the National Screening Committee using the evidence provided in this and other reports.

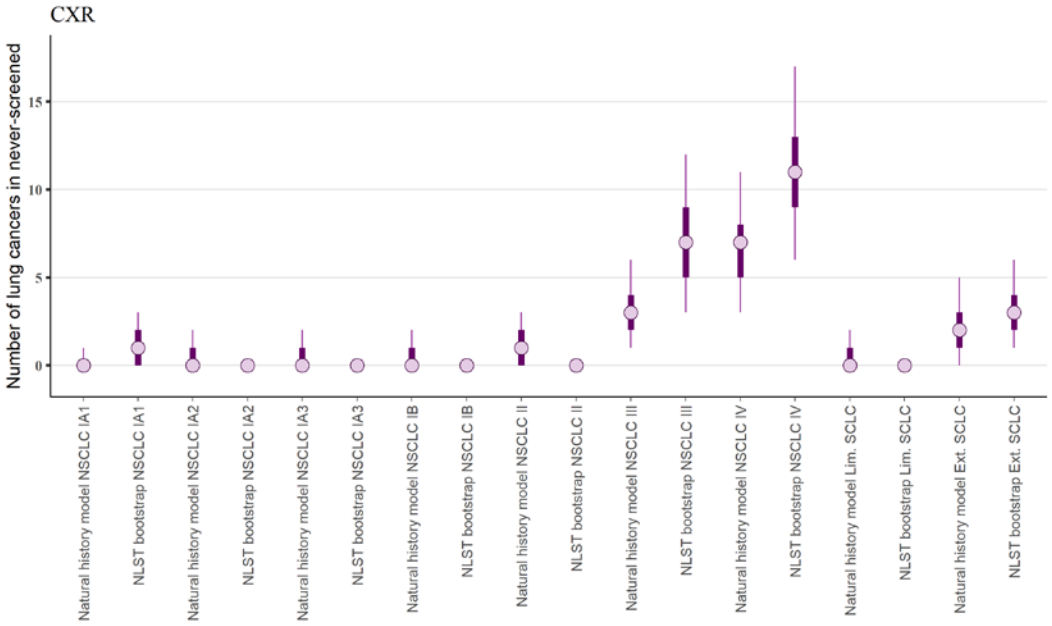
Appendix 1 — Posterior predictive distributions from the natural history model

Figure 21 Comparison of posterior predictive distribution and NLST bootstrap (a) Lung cancers in those never screened in LDCT arm; (b) Lung cancers in those never screened in CXR arm

(a)

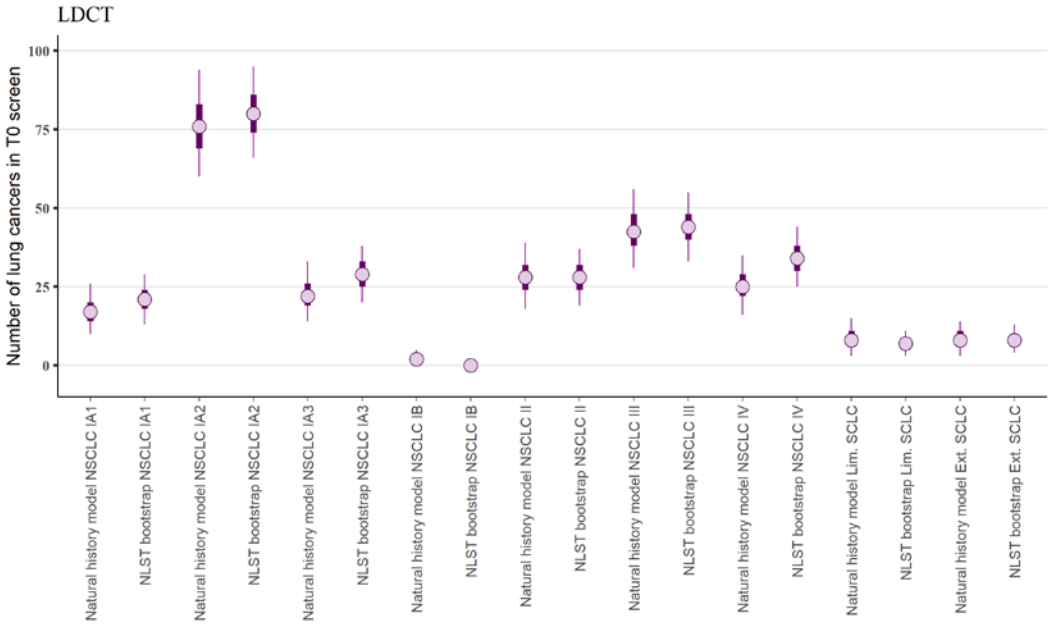


(b)

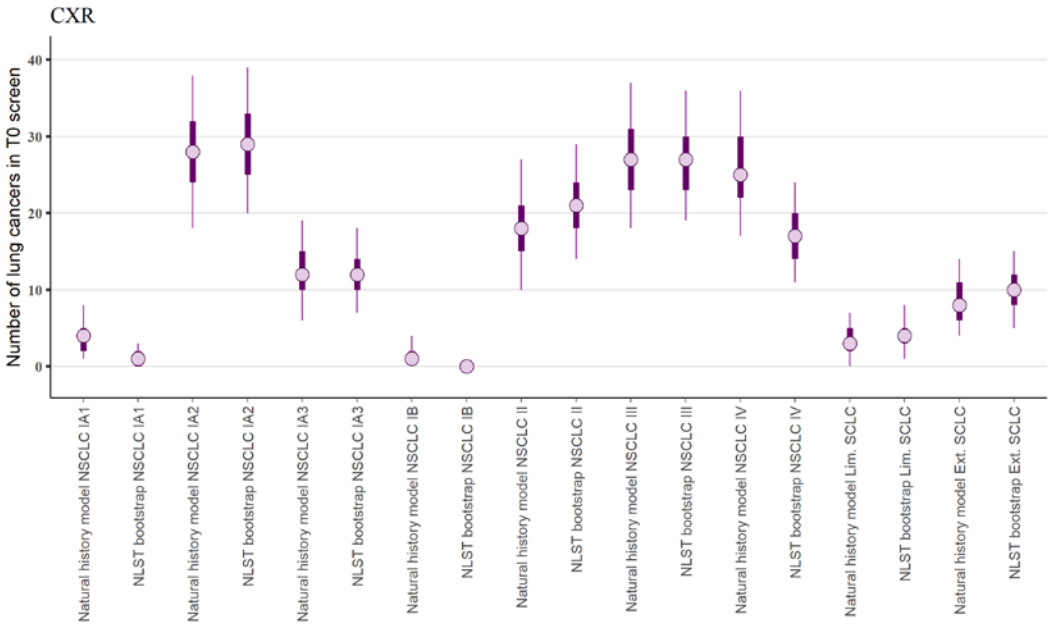


(c) Lung cancers at the baseline (T0) screen in the LDCT arm; (d) Lung cancers at the T0 screen in the CXR arm

(c)

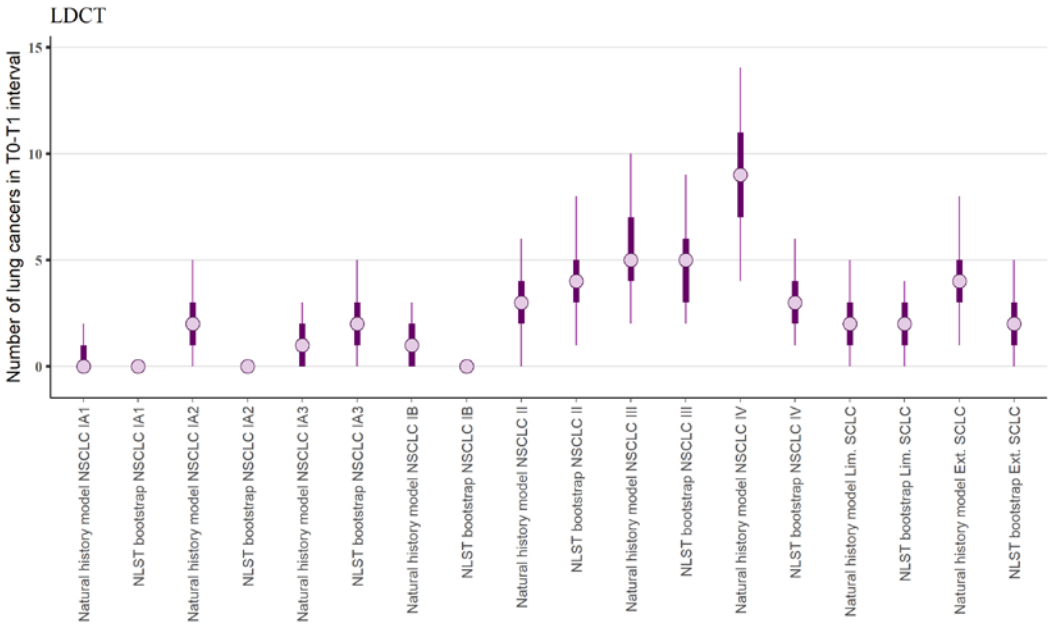


(d)

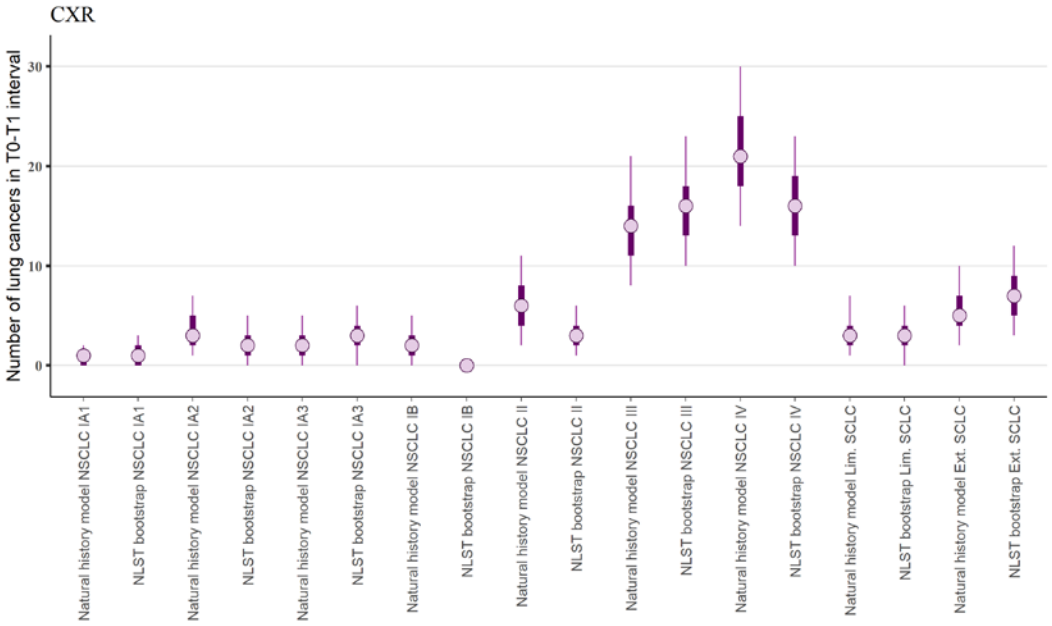


(e) Lung cancers diagnosed in the first (T0–T1) interval in the LDCT arm; (f) Lung cancers diagnosed in the T0–T1 interval in the CXR arm

(e)

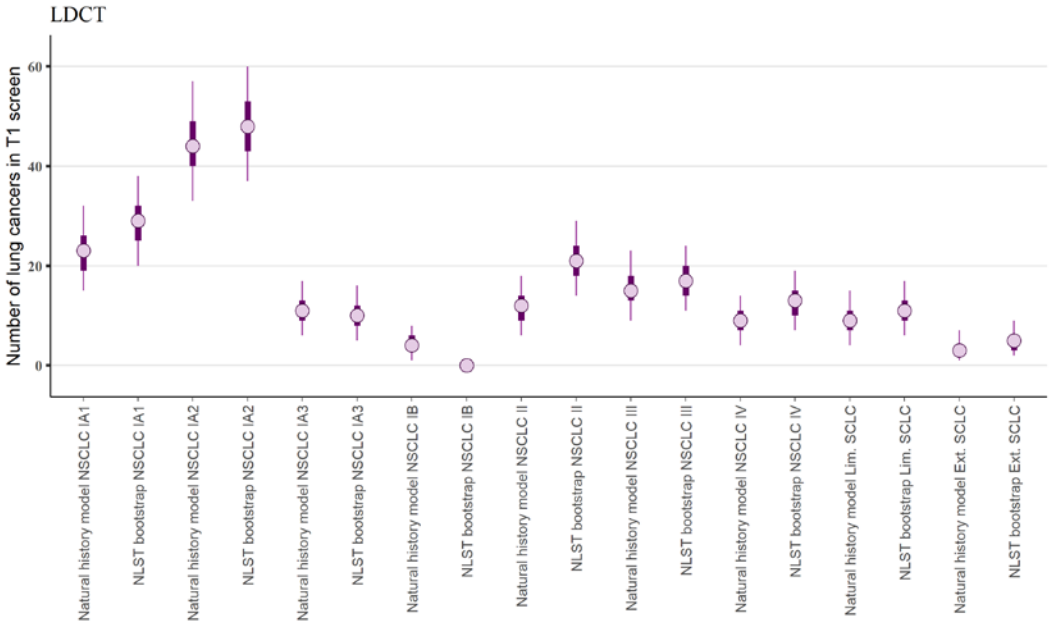


(f)

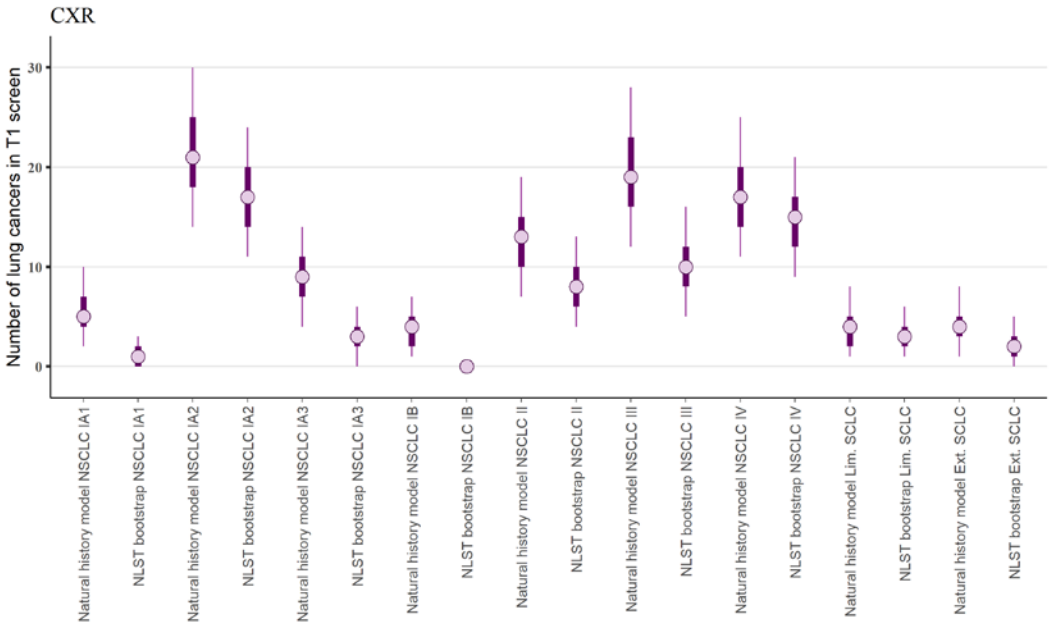


(g) Lung cancers diagnosed at the T1 screen in the LDCT arm; (h) Lung cancers diagnosed at the T1 screen in the CXR arm

(g)

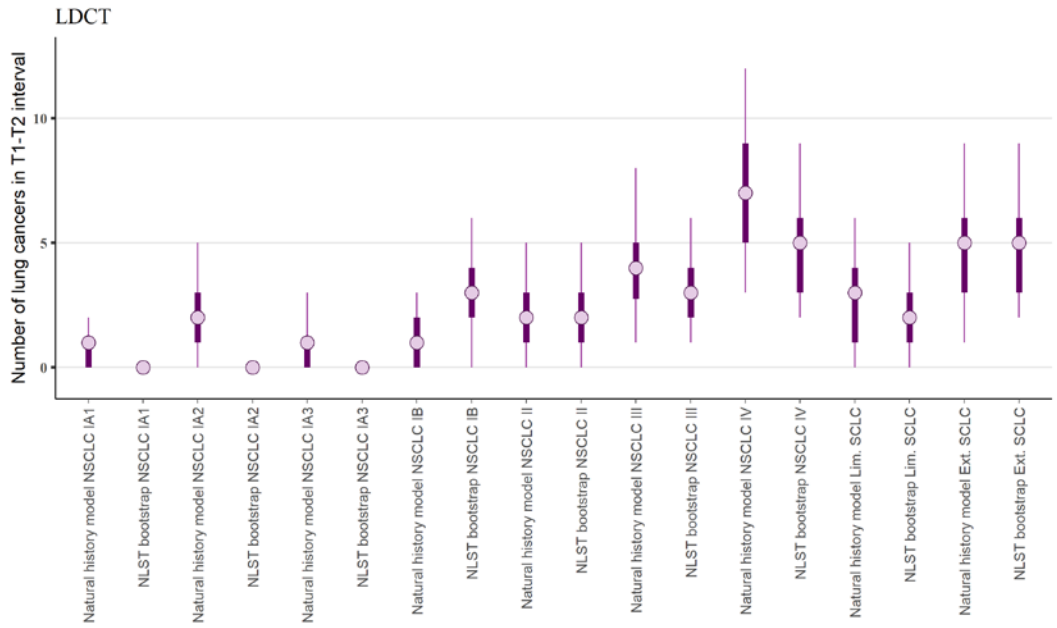


(h)

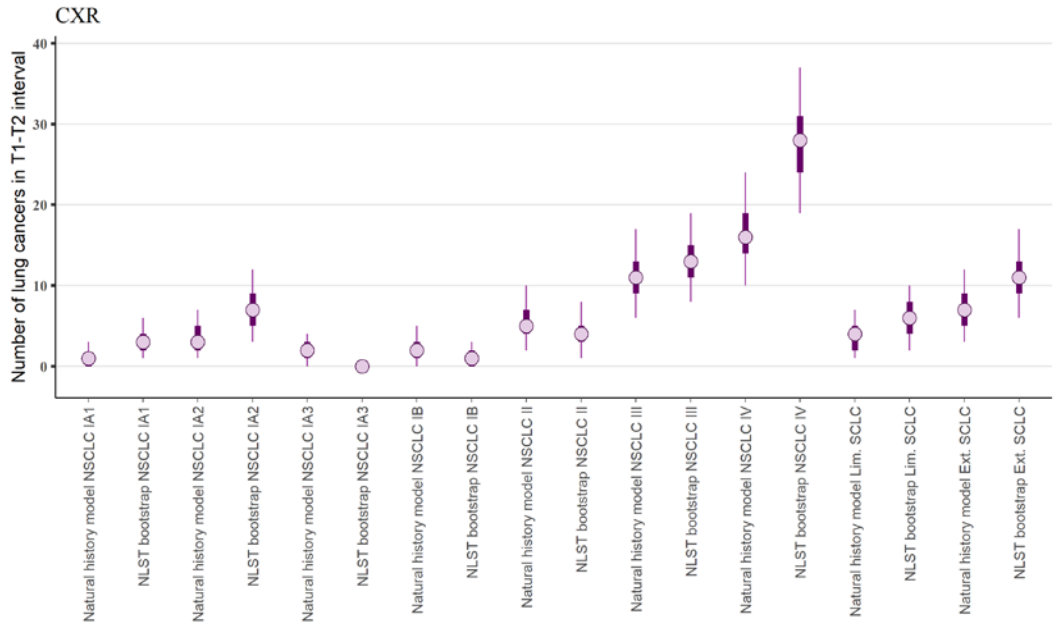


(i) Lung cancers diagnosed in the second (T1–T2) interval in the LDCT arm; (j) Lung cancers diagnosed in the second (T1–T2) interval in the CXR arm

(i)

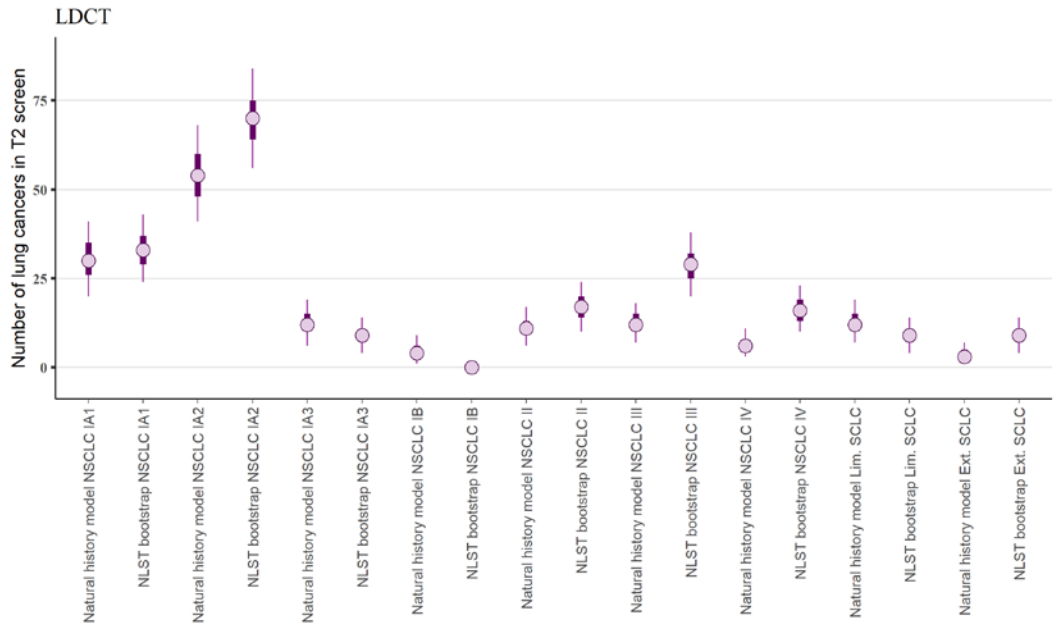


(j)

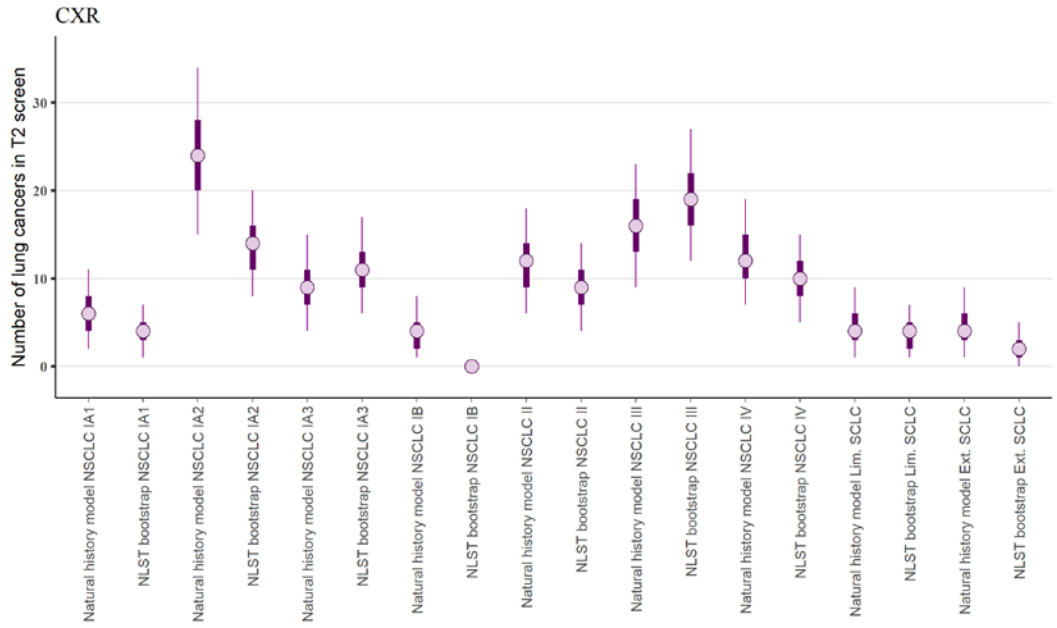


(k) Lung cancers diagnosed at the final (T2) screen in the LDCT arm; (l) Lung cancers diagnosed at the T2 screen in the CXR arm

(k)

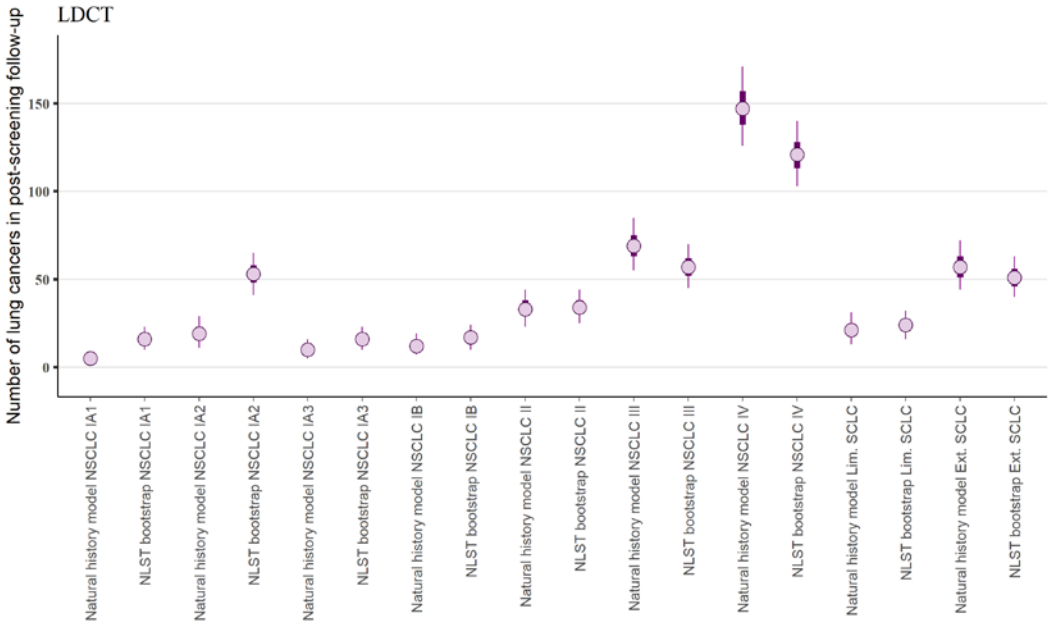


(l)

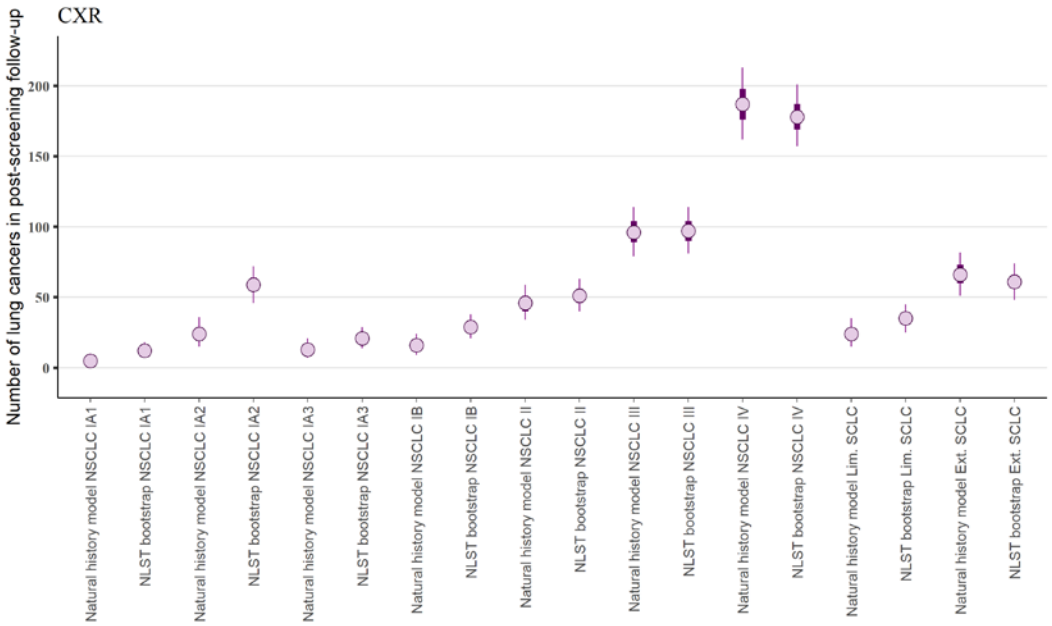


(m) Lung cancers diagnosed post-screening in the LDCT arm; (n) Lung cancers diagnosed post-screening in the CXR arm

(m)

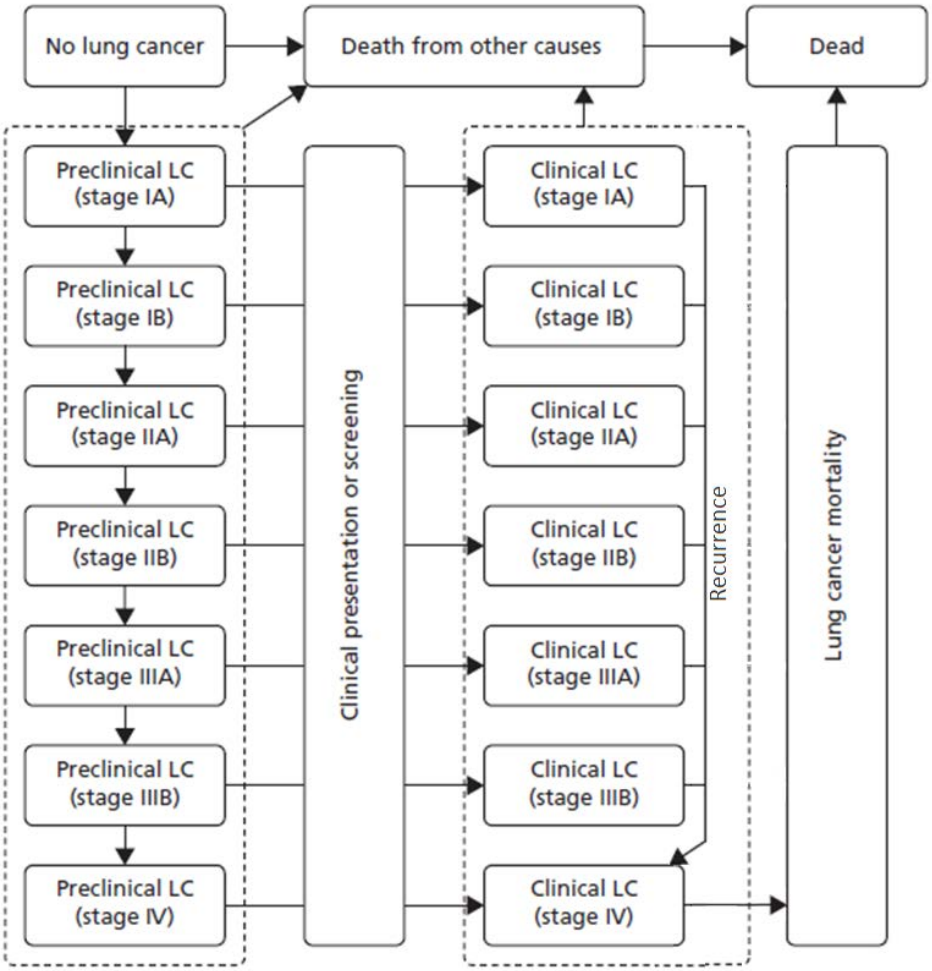


(n)



Appendix 2 – Original model structure diagram

Figure 22 Model diagram from the original report(5)



Appendix 3 — Estimation of lung cancer survival

Figure 23 Survival for a 70-year old individual diagnosed with stage II NSCLC by route of detection

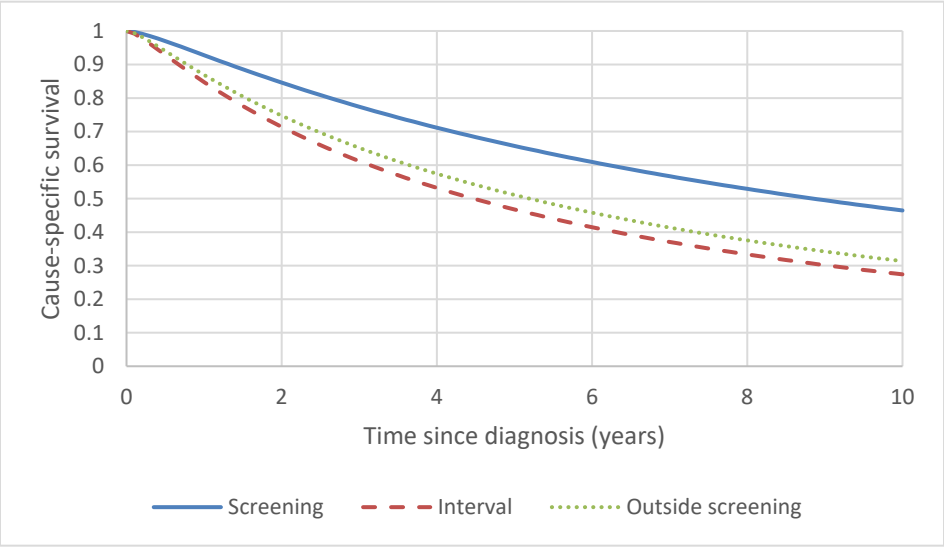
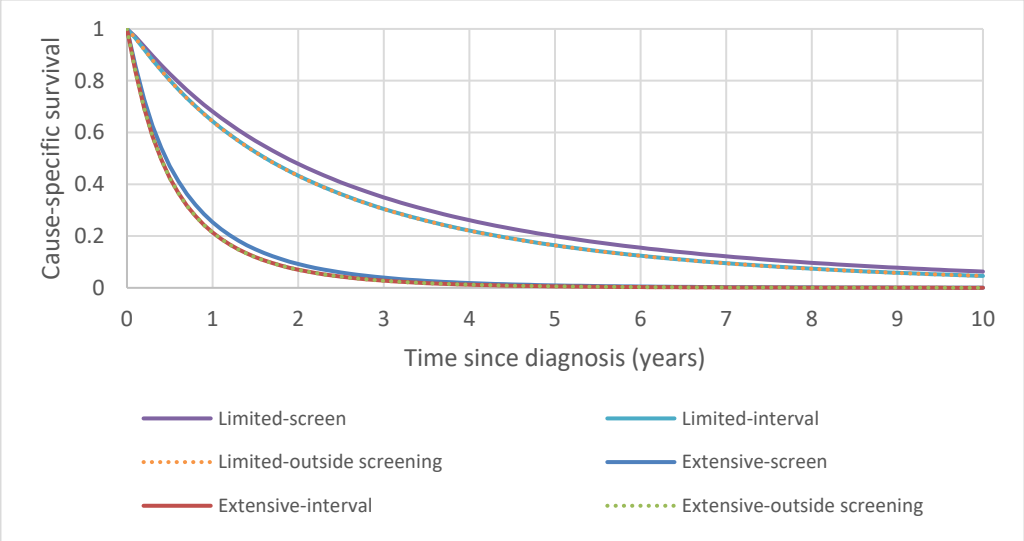
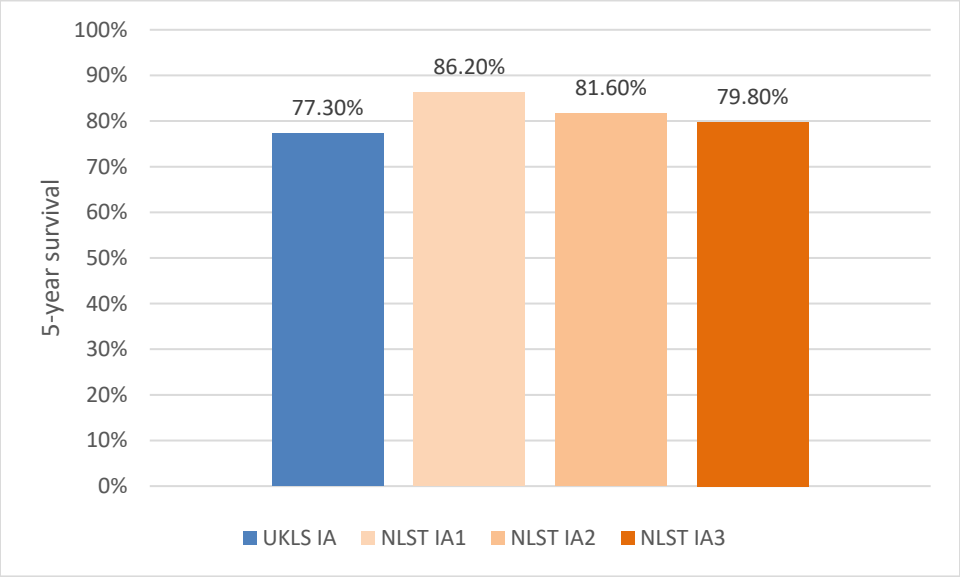


Figure 24 Survival for a 70-year old individual diagnosed with SCLC by route of detection and stage at diagnosis



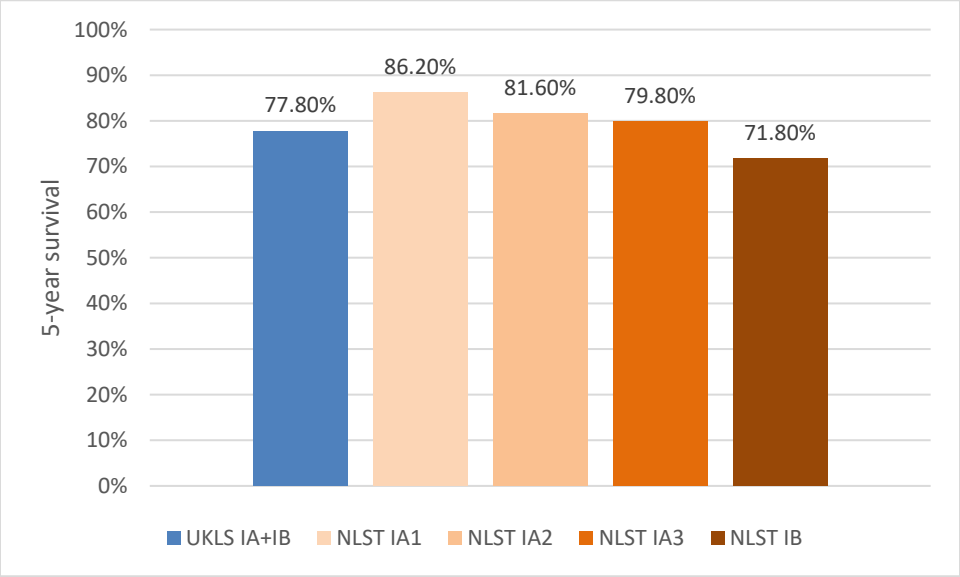
Comparison of NLST estimated survival with UK empirical evidence

Figure 25 Screen-detected 5-yr survival for NSCLC stage IA from UKLS and NLST analysis



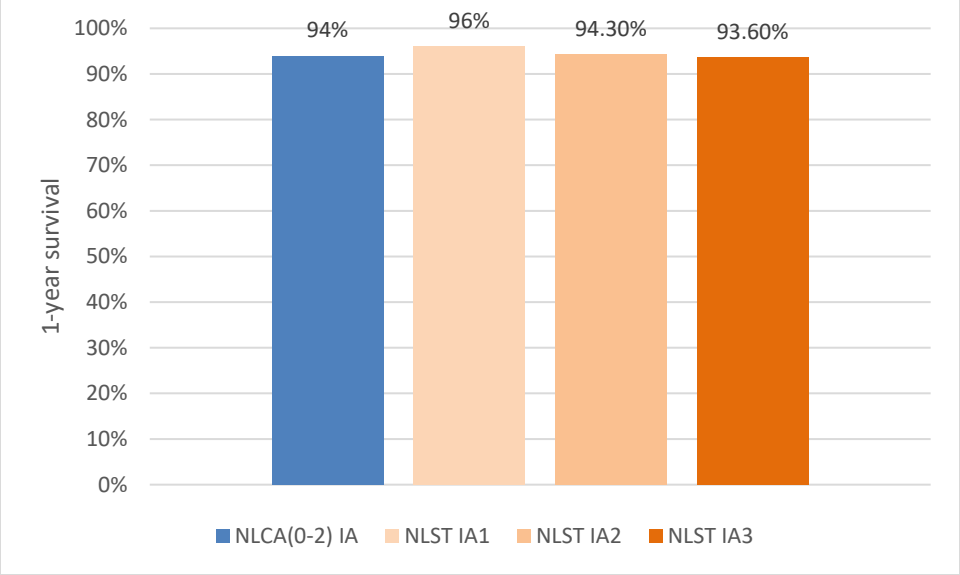
NSCLC, non-small cell lung cancer; NLST, National Lung Screen Trial; UKLS, United Kingdom Lung cancer Screening

Figure 26 Screen-detected 5-yr survival for NSCLC stage IA and IB from UKLS and NLST analysis



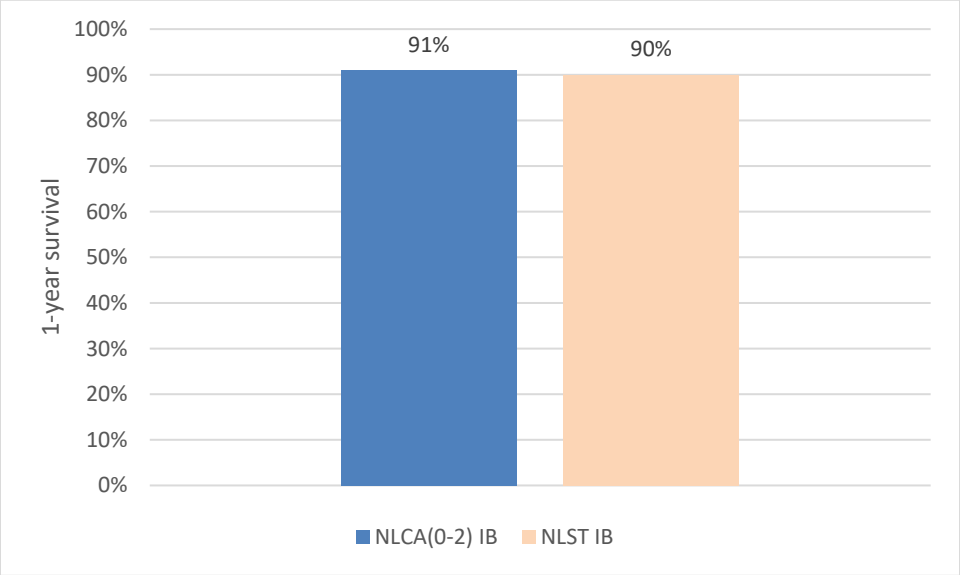
NSCLC, non-small cell lung cancer; NLST, National Lung Screen Trial; UKLS, United Kingdom Lung cancer Screening

Figure 27 Clinically-detected 1-yr survival for stage IA NLCA and NLST analysis



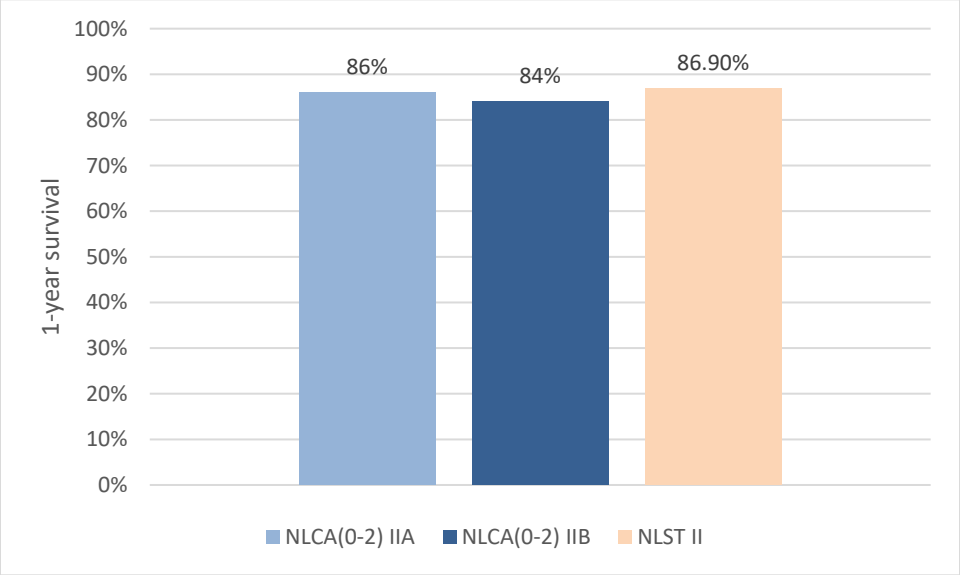
NSCLC, non-small cell lung cancer; NLCA, National Lung Cancer Audit; NLST, National Lung Screen Trial

Figure 28 Clinically-detected 1-yr survival for stage IB NLCA and NLST analysis



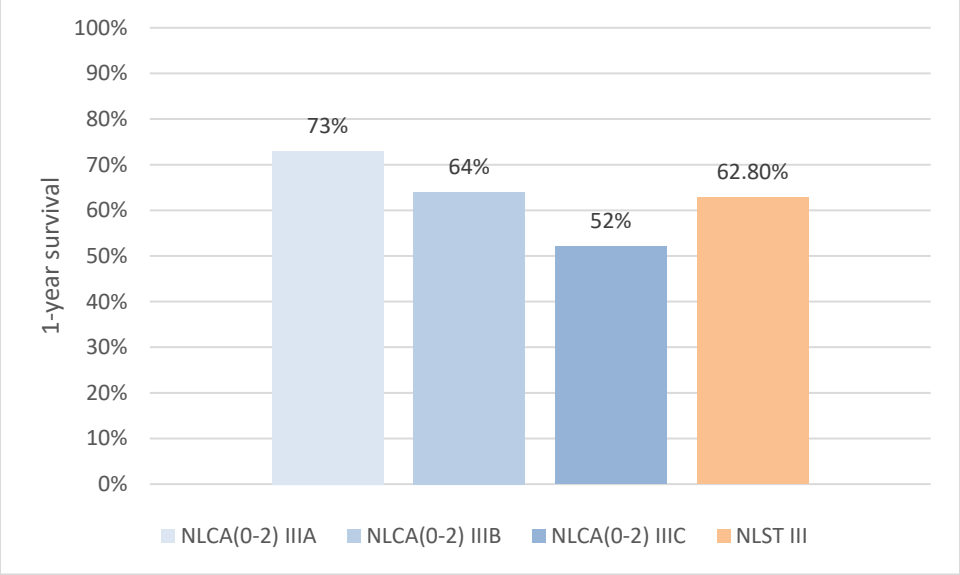
NSCLC, non-small cell lung cancer; NLCA, National Lung Cancer Audit; NLST, National Lung Screen Trial

Figure 29 Clinically-detected 1-yr survival for NSCLC stage II NLCA and NLST analysis



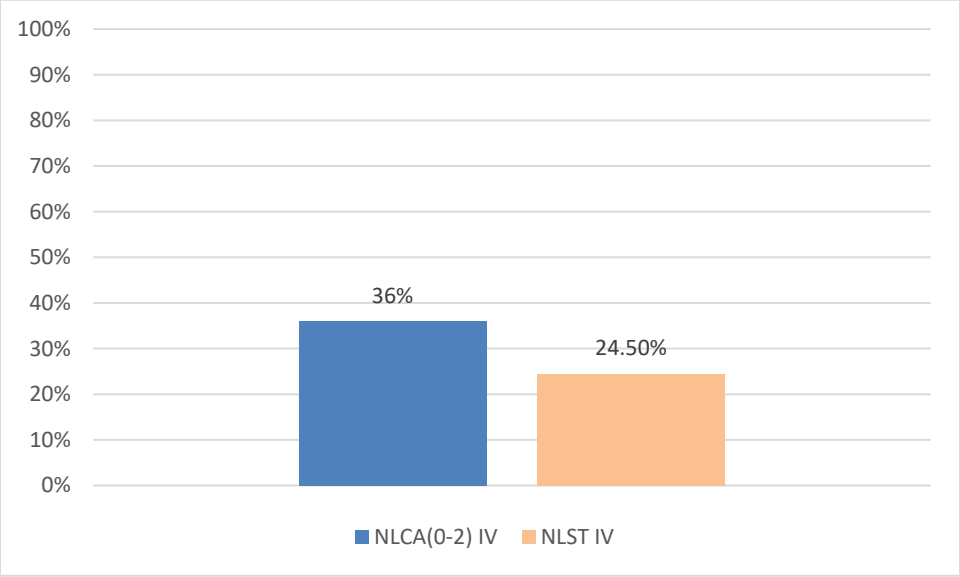
NSCLC, non-small cell lung cancer; NLCA, National Lung Cancer Audit; NLST, National Lung Screen Trial

Figure 30 Clinically-detected 1-yr survival for stage III NLCA and NLST analysis



NSCLC, non-small cell lung cancer; NLCA, National Lung Cancer Audit; NLST, National Lung Screen Trial

Figure 31 Clinically-detected 1-yr survival for NSCLC stage IV NLCA and NLST analysis



NSCLC, non-small cell lung cancer; NLCA, National Lung Cancer Audit; NLST, National Lung Screen Trial

References

1. Peters JL, Snowsill TM, Griffin E, Robinson S, Hyde CJ. Variation in Model-Based Economic Evaluations of Low-Dose Computed Tomography Screening for Lung Cancer: A Methodological Review. Value in Health. 2022.
2. Whynes DK. Could CT screening for lung cancer ever be cost effective in the United Kingdom? Cost effectiveness and resource allocation : C/E. 2008;6:5.
3. 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 technology assessment (Winchester, England). 2016;20(40):1-146.
4. 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.
5. 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 technology assessment (Winchester, England). 2018;22(69):1-276.
6. ten Haaf K, Tammemägi 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 medicine. 2017;14(2):e1002225.
7. 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 (Amsterdam, Netherlands). 2018;121:61-9.
8. 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(11):1841-8.
9. 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. Annals of Internal Medicine. 2019;171(11):796-804.
10. Tramontano AC, Schrag DL, Malin JK, Miller MC, Weeks JC, Swan JS, et al. Catalog and comparison of societal preferences (utilities) for lung cancer health states: results from the Cancer Care Outcomes Research and Surveillance (CanCORS) study. Medical decision making : an international journal of the Society for Medical Decision Making. 2015;35(3):371-87.
11. 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(3):pkz035.
12. Tammemägi MC, Katki HA, Hocking WG, Church TR, Caporaso N, Kvale PA, et al. Selection criteria for lung-cancer screening. The New England journal of medicine. 2013;368(8):728-36.
13. 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(4):405-9.

14. Church TR, Black WC, Aberle DR, Berg CD, Clingan KL, Duan F, et al. Results of initial low-dose computed tomographic screening for lung cancer. The New England journal of medicine. 2013;368(21):1980-91.
15. Moyer VA. Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2014;160(5):330-8.
16. Karnon J, Goyder E, Tappenden P, McPhie S, Towers I, Brazier J, et al. A review and critique of modelling in prioritising and designing screening programmes. Health technology assessment (Winchester, England). 2007;11(52):iii-iv, ix-xi, 1-145.
17. Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. The New England journal of medicine. 2011;365(5):395-409.
18. Karlsson A, Jauhiainen A, Gulati R, Eklund M, Grönberg H, Etzioni R, et al. A natural history model for planning prostate cancer testing: Calibration and validation using Swedish registry data. PLoS One. 2019;14(2):e0211918.
19. Kim JJ, Kuntz KM, Stout NK, Mahmud S, Villa LL, Franco EL, et al. Multiparameter calibration of a natural history model of cervical cancer. Am J Epidemiol. 2007;166(2):137-50.
20. Whyte S, Walsh C, Chilcott J. Bayesian calibration of a natural history model with application to a population model for colorectal cancer. Medical decision making : an international journal of the Society for Medical Decision Making. 2011;31(4):625-41.
21. Stan Development Team. Stan Modeling Language Users Guide and Reference Manual, v2.26. 2021.
22. Gray EP, Teare MD, Stevens J, Archer R. Risk Prediction Models for Lung Cancer: A Systematic Review. Clin Lung Cancer. 2016;17(2):95-106.
23. Ten Haaf K, Jeon J, Tammemägi MC, Han SS, Kong CY, Plevritis SK, et al. Risk prediction models for selection of lung cancer screening candidates: A retrospective validation study. PLoS medicine. 2017;14(4):e1002277.
24. Bradley SH, Abraham S, Callister ME, Grice A, Hamilton WT, Lopez RR, et al. Sensitivity of chest X-ray for detecting lung cancer in people presenting with symptoms: a systematic review. Br J Gen Pract. 2019;69(689):e827-e35.
25. 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(1):154-61.
26. Heuvelmans MA, Vliegenthart R, de Koning HJ, Groen HJM, van Putten M, Yousaf-Khan U, et al. Quantification of growth patterns of screen-detected lung cancers: The NELSON study. Lung cancer (Amsterdam, Netherlands). 2017;108:48-54.
27. Betancourt M. A conceptual introduction to Hamiltonian Monte Carlo. arXiv preprint arXiv:170102434. 2017.
28. Claxton K, Paulden M, Gravelle H, Brouwer W, Culyer AJ. Discounting and decision making in the economic evaluation of health-care technologies. Health Econ. 2011;20(1):2-15.
29. Paulden M. Calculating and Interpreting ICERs and Net Benefit. Pharmacoeconomics. 2020;38(8):785-807.
30. Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WE, et al. The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer. J Thorac Oncol. 2016;11(1):39-51.

31. Schabath MB, Massion PP, Thompson ZJ, Eschrich SA, Balagurunathan Y, Goldof D, et al. Differences in Patient Outcomes of Prevalence, Interval, and Screen-Detected Lung Cancers in the CT Arm of the National Lung Screening Trial. PLoS One. 2016;11(8):e0159880.
32. Khakwani A, Rich AL, Tata LJ, Powell HA, Stanley RA, Baldwin DR, et al. Small-cell lung cancer in England: trends in survival and chemotherapy using the National Lung Cancer Audit. PLoS One. 2014;9(2):e89426.
33. Office for National Statistics. National life tables - life expectancy in the UK: 2017 to 2019. 2021.
34. Cassidy A, Myles JP, van Tongeren M, Page RD, Liloglou T, Duffy SW, et al. The LLP risk model: an individual risk prediction model for lung cancer. British journal of cancer. 2008;98(2):270-6.
35. NatCen Social Research and UCL. Health Survey for England 2019. Adults health-related behaviours NHS Digital2020 [Available from: https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/2019.
36. Lebrett MB, Balata H, Evison M, Colligan D, Duerden R, Elton P, et al. Analysis of lung cancer risk model (PLCO(M2012) and LLP(v2)) performance in a community-based lung cancer screening programme. Thorax. 2020;75(8):661-8.
37. Warkentin MT, Lam S, Hung RJ. Determinants of impaired lung function and lung cancer prediction among never-smokers in the UK Biobank cohort. EBioMedicine. 2019;47:58-64.
38. National Cancer Registration and Analysis Service (NCRAS). England Cancer Prevalence Statistics, 2019 NHS Digital2022 [Available from: https://www.cancerdata.nhs.uk/prevalence.
39. Brown KF, Rumgay H, Dunlop C, Ryan M, Quartly F, Cox A, et al. The fraction of cancer attributable to modifiable risk factors in England, Wales, Scotland, Northern Ireland, and the United Kingdom in 2015. British journal of cancer. 2018;118(8):1130-41.
40. NatCen Social Research and UCL. Health Survey for England 2010. Respiratory health NHS Digital2011 [Available from: https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/health-survey-for-england-2010-respiratory-health.
41. Office for National Statistics. Adult smoking habits in Great Britain Office for National Statistics,2021 [Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/drugusealcoholandsmoking/datasets/adultsmokinghabitsingreatbritain.
42. Ghimire B, Maroni R, Vulkan D, Shah Z, Gaynor E, Timoney M, et al. Evaluation of a health service adopting proactive approach to reduce high risk of lung cancer: The Liverpool Healthy Lung Programme. Lung cancer (Amsterdam, Netherlands). 2019;134:66-71.
43. 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 (Amsterdam, Netherlands). 2020;148:12-9.
44. Quaife SL, Ruparel M, Dickson JL, Beeken RJ, McEwen A, Baldwin DR, et al. Lung Screen Uptake Trial (LSUT): Randomized Controlled Clinical Trial Testing Targeted Invitation Materials. American journal of respiratory and critical care medicine. 2020;201(8):965-75.
45. Lung cancer screening pathways. Public Health England,; 2021.
46. Crosbie PA, Gabe R, Simmonds I, Kennedy M, Rogerson S, Ahmed N, et al. Yorkshire Lung Screening Trial (YLST): protocol for a randomised controlled trial to evaluate invitation to community-based low-dose CT screening for lung cancer versus usual care in a targeted population at risk. BMJ open. 2020;10(9):e037075.

47. U.S. National Library of Medicine. The SUMMIT Study: A Cancer Screening Study [Available from: https://www.clinicaltrials.gov/ct2/show/NCT03934866.
48. 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(7):700-4.
49. Horeweg N, Scholten ET, de Jong PA, van der Aalst CM, Weenink C, Lammers JW, et al. Detection of lung cancer through low-dose CT screening (NELSON): a prespecified analysis of screening test performance and interval cancers. The Lancet Oncology. 2014;15(12):1342-50.
50. Field JK, Vulkan D, Davies MPA, Baldwin DR, Brain KE, Devaraj A, et al. Lung cancer mortality reduction by LDCT screening: UKLS randomised trial results and international meta-analysis. The Lancet regional health Europe. 2021;10:100179.
51. Callister ME, Baldwin DR, Akram AR, Barnard S, Cane P, Draffan J, et al. British Thoracic Society guidelines for the investigation and management of pulmonary nodules. Thorax. 2015;70 Suppl 2:ii1-ii54.
52. Blom EF--H, K. T.-//-de Koning, H. J. Systematic Review and Meta-Analysis of Community- and Choice-Based Health State Utility Values for Lung Cancer. Pharmacoeconomics. 2020;38(11):1187-200.
53. Pourrahmat M-M, Kim A, Kansal AR, Hux M, Pushkarna D, Fazeli MS, et al. Health state utility values by cancer stage: a systematic literature review. The European journal of health economics : HEPAC : health economics in prevention and care. 2021;22(8):1275-88.
54. Yang SC--K, C. W.-//-Lai, W. W.-//-Lin, C. C.-//-Su, W. C.-//-Chang, S. M.-//-Wang, J. D. Dynamic Changes of Health Utility in Lung Cancer Patients Receiving Different Treatments: A 7-Year Follow-up. J Thorac Oncol. 2019;14(11):1892-900.
55. Zeng X--S, M.-//-Liu, B.-//-Yang, H.-//-Liu, R.-//-Tan, R. L. Y.-//-Xu, J.-//-Zheng, E.-//-Yang, J.-//-Liu, C.-//-Huang, W.-//-Yu, H.-//-Luo, N. Measurement Properties of the EQ-5D-5L and EQ-5D-3L in Six Commonly Diagnosed Cancers. Patient. 2020.
56. Liu L, Wei Y, Teng Y, Yan J, Li F, Chen Y. Health-Related Quality of Life and Utility Scores of Lung Cancer Patients Treated with Traditional Chinese Medicine in China. Patient preference and adherence. 2022;16:297-306.
57. Kuehne N, Hueniken K, Xu M, Shakik S, Vedadi A, Pinto D, et al. Longitudinal Assessment of Health Utility Scores, Symptoms and Toxicities in Patients with Small Cell Lung Cancer Using Real World Data. Clinical lung cancer. 2022;23(2):e154-e64.
58. Taghizadeh N--T, A.-//-Cressman, S.-//-Peacock, S.-//-McWilliams, A. M.-//-MacEachern, P.-//-Johnston, M. R.-//-Goffin, J.-//-Goss, G.-//-Nicholas, G.-//-Martel, S.-//-Laberge, F.-//-Bhatia, R.-//-Liu, G.-//-Schmidt, H.-//-Atkar-Khattra, S.-//-Tsao, M. S.-//-Tammemagi, M. C.-//-Lam, S. C.-//-Pan-Canadian Early Lung Cancer Study, Group. Health-related quality of life and anxiety in the PAN-CAN lung cancer screening cohort. BMJ open. 2019;9(1):e024719.
59. Li L--S, J. L. H.-//-Mandrik, O. Disutility associated with cancer screening programs: A systematic review. PLoS ONE [Electronic Resource]. 2019;14(7):e0220148.
60. Craig R, Mindell J. Health Survey for England 2014: Health, Social Care and Lifestyles. Summary of Key Findings. NHS Digial2014.
61. Sturza J. A review and meta-analysis of utility values for lung cancer. Medical decision making : an international journal of the Society for Medical Decision Making. 2010;30(6):685-93.
62. van den Bergh KA, Essink-Bot ML, Borsboom GJ, Th Scholten E, Prokop M, de Koning HJ, et al. Short-term health-related quality of life consequences in a lung cancer CT screening trial (NELSON). British journal of cancer. 2010;102(1):27-34.

63. Mazzone PJ, Obuchowski N, Fu AZ, Phillips M, Meziane M. Quality of life and healthcare use in a randomized controlled lung cancer screening study. Annals of the American Thoracic Society. 2013;10(4):324-9.
64. GOV.UK Digital Marketplace. Health Information Consulting Ltd, Patient Provider Index 2022 [Available from: https://www.digitalmarketplace.service.gov.uk/g-cloud/services/844828413177937.
65. NatCen Social Research UCL. Health survey for England 2015. Trend tables commentary. NHS Digital; 2016.
66. GOV.UK Digital Marketplace. NEC Software Solutions UK Limited, NEC Abdominal Aortic Aneurysm (AAA) SMaRT Version 9.5.1 2022 [Available from: https://www.digitalmarketplace.service.gov.uk/g-cloud/services/486650661491393.
67. NHS. 2019/20 National Cost Collection Data Publication. 2020.
68. Harpaz SB, Weber MF, Wade S, Ngo PJ, Vaneckova P, Sarich PEA, et al. Updated cost-effectiveness analysis of lung cancer screening for Australia, capturing differences in the health economic impact of NELSON and NLST outcomes. British journal of cancer. 2022.
69. Department of Health and Social Care. NHS Reference Costs 2015 to 2016. London: Department of Health and Social Care; 2016.
70. Incisive Health. Saving lives, averting costs: An analysis of the financial implications of achieving earlier diagnosis of colorectal, lung and ovarian cancer. Cancer Research UK; 2014.
71. Goldstraw P, Crowley J, Chansky K, Giroux DJ, Groome PA, Rami-Porta R, et al. The IASLC Lung Cancer Staging Project: proposals for the revision of the TNM stage groupings in the forthcoming (seventh) edition of the TNM Classification of malignant tumours. J Thorac Oncol. 2007;2(8):706-14.
72. Kennedy MP, Hall PS, Callister ME. Factors affecting hospital costs in lung cancer patients in the United Kingdom. Lung cancer (Amsterdam, Netherlands). 2016;97:8-14.
73. 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. Journal of medical economics. 2015;18(7):525-32.