

# Imaging-based biological age estimation predicts competing mortality risk in lung cancer screening

**Daryl Cheng**

daryl.cheng@ucl.ac.uk

University College London

**John McCabe**

UCL Hawkes Institute and Department of Computer Science, University College London, UK

**Tanya Patrick**

Lungs for Living Research Centre, University College London, UK

**Robert Chapman**

UCL Hawkes Institute and Department of Computer Science, University College London, UK

**Shanshan Wang**

UCL Hawkes Institute and Department of Computer Science, University College London, UK

**Tianqi Yang**

UCL Hawkes Institute and Department of Computer Science, University College London, UK

**Adam Szmul**

UCL Hawkes Institute and Department of Computer Science, University College London, UK

**Mehran Azimbagirad**

UCL Hawkes Institute and Department of Computer Science, University College London, UK

**Shahab Aslani**

UCL Hawkes Institute and Department of Computer Science, University College London, UK

**Monica Mullin**

Division of Respiratory Medicine, Cumming School of Medicine, University of Calgary, Canada

**Sindhu Bhaarrati Naidu**

Lungs for Living Research Centre, University College London, UK

**Chuen Ryan Khaw**

Lungs for Living Research Centre, University College London, UK

**Alfred So**

UCL Hawkes Institute and Department of Computer Science, University College London, UK

**Neal Navani**

University College London Hospital

**Arjun Nair**

Department of Radiology, University College London Hospitals NHS Trust, UK

**John Hurst**

University College London

**Andre Altmann**

The UCL Hawkes Institute, University College London, London, UK

**Adam Pennycuik**

Lungs for Living Research Centre, University College London, UK

**Daniel Alexander**

Centre for Medical Image Computing, Department of Computer Science, University College London

**Neil Oxtoby**

UCL Hawkes Institute and Department of Computer Science, University College London, UK

**Alexandra Young**

UCL Hawkes Institute and Department of Computer Science, University College London, UK

**Matthew Sperrin**

Division of Informatics, Imaging & Data Sciences, University of Manchester, UK

**Sam Janes**

Lungs for Living Research Centre, University College London, UK

**Joseph Jacob**

University College London

---

**Article**

**Keywords:**

**Posted Date:** August 10th, 2026

**DOI:** <https://doi.org/10.21203/rs.3.rs-9073407/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

**Version of Record:** A version of this preprint was published at Scientific Reports on October 15th, 2026. See the published version at <https://doi.org/10.1038/s41598-026-75832-2>.

# Imaging-based biological age estimation predicts competing mortality risk in lung cancer screening

Daryl O. Cheng <sup>1,2, ‡</sup>

John McCabe <sup>1,2</sup>

Tanya Patrick <sup>2</sup>

Robert Chapman <sup>1</sup>

Shanshan Wang <sup>1</sup>

Tianqi Yang <sup>1</sup>

Adam Szmul <sup>1</sup>

Mehran Azimbagirad <sup>1</sup>

Shahab Aslani <sup>1,3</sup>

Monica L. Mullin <sup>4</sup>

Sindhu Bhaarrati Naidu <sup>2</sup>

Chuen Ryan Khaw <sup>2</sup>

Alfred CP So <sup>1</sup>

SUMMIT consortium <sup>\*</sup>

Neal Navani <sup>1</sup>

Arjun Nair <sup>5</sup>

John R Hurst <sup>3</sup>

Andre Altmann <sup>6</sup>

Adam Pennycuik <sup>2</sup>

Daniel C. Alexander <sup>1</sup>

Neil P. Oxtoby <sup>1</sup>

Alexandra L. Young <sup>1</sup>

Matthew Sperrin <sup>7</sup>

Sam M. Janes <sup>2,†</sup>

Joseph Jacob <sup>1,†</sup>

1: UCL Hawkes Institute and Department of Computer Science, University College London, UK

2: Lungs for Living Research Centre, University College London, UK

3: UCL Respiratory, University College London, UK

4: Division of Respiratory Medicine, Cumming School of Medicine, University of Calgary, Canada

5: Department of Radiology, University College London Hospitals NHS Trust, UK

6: UCL Hawkes Institute and Department of Medical Physics and Biomedical Engineering, University College London, UK

7: Division of Informatics, Imaging & Data Sciences, University of Manchester, UK

\*: listed in acknowledgements

†: represents joint senior authorship

‡: Corresponding author: daryl.cheng@ucl.ac.uk

Figure count: 6

Table count: 2

This article has accompanying Supplementary Materials.

## Abstract

**Background:** Lung cancer screening (LCS) participants have a high competing risk of non-lung cancer related death, which limits screening benefit. We hypothesized that imaging-based biological age can identify individuals at higher risk of non-lung cancer mortality.

**Methods:** Low dose computed tomography scans from a large LCS trial were retrospectively analyzed. Bone density loss, muscular fat infiltration, vascular calcification, and visceral fat mass were calculated using segmentations from a nnU-Net based deep learning segmentation model. Each participant's age gap, defined as the difference between their estimated biological age and chronological age, was estimated with disease course mapping using a Bayesian mixed-effects model. Disease course map construction used all available longitudinal scan data, whilst individual biological age prediction used baseline scan data only. The performance of age gap as a parameter in competing risk models for non-lung cancer related death versus lung cancer diagnosis was evaluated, following the TRIPOD+AI reporting guideline.

**Findings:** We included 29,745 scans from 12,478 participants in the final analysis. Median follow-up was 5 years, representing 58,284 person-years at risk. Multivariable survival models of competing risks were constructed, including the estimated age gap. The subdistribution hazard ratio (sHR) for each decade of age gap for non-lung cancer related death was higher (sHR=2.0, 95% CI 1.8–2.3), compared to lung cancer diagnosis (sHR=1.3, 95% CI 1.2–1.5). Addition of age gap improved model concordance of non-lung cancer death compared to a model with clinical variables alone (pooled C-index=0.69 vs 0.73, difference +0.04, 95% CI 0.03–0.04,  $p<0.001$ ).

**Interpretation:** Biological age estimation using an imaging-based disease progression model improves prediction of non-lung cancer related death in lung cancer screening.

**Funding:** None

## Introduction

Decades of evidence have shown that lung cancer screening (LCS) reduces lung cancer-related mortality<sup>1,2</sup>. However, improvement in overall survival is less clear, with large LCS trials reporting no overall mortality benefit at 10 years<sup>2</sup>. Selecting a population at high risk of lung cancer also selects participants at high risk of death from a multitude of other causes, such as cardiovascular disease, chronic obstructive pulmonary disease (COPD), and other cancers. Neglecting to account for competing mortality risks exacerbating overdiagnosis in screening. Improving overall survival in this population requires disentangling the competing risks of lung cancer and other comorbidity and frailty.

Striking the balance of these risks has been the topic of active research. Approaches investigating individual response to screening based on frailty such as the Electronic Frailty Index<sup>3</sup> using health record data, or participant selection based on predicted risk of non-lung cancer death<sup>4</sup>, or expected gains in life-expectancy<sup>5</sup> are promising methods to enable more equitable and effective screening, with patient centered, shared decision-making. The use of imaging in lung cancer screening has often focused on refining lung cancer risk, such as from nodule malignancy prediction<sup>6</sup>. Imaging measures of emphysema and cardiovascular disease may aid in the refinement of risk of non-cancer related death<sup>7</sup>. However, prediction of non-lung cancer related death using clinically understandable metrics in competing risks time-to-event models is less well studied.

Improving prediction of non-lung cancer related death in screening may help identify individuals whose competing mortality risk limits potential screening benefit. We hypothesized that estimates of biological age can predict the risk of non-lung cancer related death. We therefore proposed the construction of a disease course map based on the evolution of imaging biomarkers to construct a single metric of imaging-derived biological age and explored its performance in predicting both cancer diagnosis and non-lung cancer death.

## Methods

### Study population

The SUMMIT Study was a longitudinal, prospective, cohort study investigating the deliverability of low dose computed tomography (LDCT) screening for lung cancer in a high-risk population in London, United Kingdom (NCT03934866, Registration date 26 April 2019, ClinicalTrials.gov). Current or former smokers aged 55 to 77 were enrolled from April 2019 to March 2020. Ethical approval for the SUMMIT Study and ongoing analyses was obtained from an NHS Research Ethics Committee (17/LO/2004) and the NHS Health Research Authority's Confidentiality Advisory Group (18/CAG/0054), and written informed consent was obtained from all participants. Clinical data collected included age, self-reported sex assigned at birth, self-reported ethnicity, smoking status, smoking duration, body mass index (BMI), COPD diagnosis at baseline, spirometry data, and rank of Index of Multiple Deprivation (IMD).

### Imaging feature extraction and normalization

LDCT scans were acquired in one continuous craniocaudal acquisition at full inspiration using GE Revolution multi-slice scanners. Scans were non-contrast at 0.625mm slice thickness. All scans were reviewed by specialist thoracic radiologists as part of the lung cancer screening workflow. An in-house nnU-Net based deep learning segmentation algorithm was adapted from previously published methods<sup>8</sup>. These segmentations were then used to generate the features used for disease course map construction (described below). The metrics were of (1) erector spinae muscle density; (2) trabecular bone density; (3) aortic calcification; and (4) pericardial fat mass. Metrics were chosen to measure the known age-related changes of muscular fat infiltration<sup>9</sup>, bone mineral density loss<sup>10</sup>, increase in vessel calcification<sup>11</sup>, and redistribution of visceral fat<sup>12</sup>. Full details of methods used for image feature extraction are given in Supplementary Materials.

## Internal-external cross validation

An Internal-external Cross Validation (IECV) workflow by geographical regions was used for disease course map model fitting and inference, competing risk modelling, and competing risk model performance evaluation. This design tests robustness to geographic and case-mix variation across a held-out cluster, with the caveat that it is performed within a single large trial with homogenous imaging protocols, and does not constitute truly distinct geographical external validation. The competing risks model performance is likewise calculated on a held-out geographical cluster, simulating case-mix variation within the same setting.<sup>13–15</sup> The study population was split into 3 geographical clusters, grouped according to International Territorial Level (Supplementary Table 4). Concordance and calibration were estimated using a pooled random effects meta-analysis. The schematic of the IECV process is represented in Figure 1 (further detail in Supplementary Materials).

## Disease course mapping

Disease progression modeling reconstructs long-term disease trajectories by positioning individuals at different stages along a common pathway. These approaches have been applied to degenerative neurological diseases<sup>16</sup> and to identify COPD progression patterns from lung CT imaging<sup>17</sup>. Whilst aging is not a disease, age-related biomarker changes follow trajectories that can be modeled using the same mathematical framework. We adapt a Bayesian mixed-effect disease course mapping model previously described in the context of Alzheimer's disease<sup>18</sup>.

All available biomarker data, whether single timepoint or multi-timepoint, were used to estimate a map of biomarker progression by mapping each subject's data to different portions of a common trajectory. By using the age at each observation, the constructed map is one of the expected trajectories of each of the imaging biomarkers by chronological age. Age gap is defined as the difference between the population mean age of onset of biomarker evolution ( $\tau_{\text{mean}}$ ) and the estimated individual age of onset of biomarker evolution ( $\tau_i$ ), with negative values indicating a younger biological age, and positive values indicating an older biological age.

## Competing risks survival analysis

Time to lung cancer diagnosis and time to non-lung cancer related death were each modelled using separate Fine-Gray subdistribution hazard models within a competing risks framework. For the non-lung cancer death model, lung cancer diagnosis was treated as the competing event, and for the lung cancer diagnosis model, non-lung cancer death was treated as the competing event. Both models estimate cumulative incidence functions (CIF) and subdistribution hazard ratios (sHR), which account for the presence of competing risks. Lung cancer diagnosis was selected rather than lung cancer death because screening benefit derives from early detection and curative treatment, and a participant who dies from competing causes before diagnosis cannot benefit from screening, regardless of cancer risk.

Follow up time was calculated from study enrollment to the outcome of lung cancer diagnosis, and to death from any other cause or censor as a competing event. Non-lung cancer related death was defined using death certificate data from the National Cancer Registration and Analysis Service, where International Classification of Disease-10 codes for lung cancer related terms did not appear in part 1a, 1b or 1c (Supplementary Materials).

The age gap metric included in competing risks models were inferred from baseline imaging in the held out IECV region using the disease course map fit to training IECV regions. Feature selection was based on clinical knowledge and existing literature for lung cancer and non-lung cancer death risk prediction in lung cancer screening<sup>4</sup>. Interactions which were deemed plausible were age-age gap, sex-smoking status, sex-smoking amount, and sex-time since stopping smoking, and included in models if  $p < 0.05$ . All multivariable models were adjusted for age, sex, BMI, maximum educational attainment, ethnicity, IMD rank quintile, diagnosis of COPD, bronchitis or emphysema at baseline, family history of lung cancer, personal history of any cancer, smoking status, smoking duration, smoking dose (pack-year equivalents), and time since quitting smoking.

## Model Comparisons

Model discrimination was assessed by concordance (C-index) at 5 years of follow-up. Calibration was reported as pooled estimates of calibration slope and intercept,

as well as visually in calibration plots. Validation in the presence of competing risks followed methods described in<sup>19,20</sup>, using pseudo-observations. Methods are described further in Supplementary Materials.

To assess whether the single integrated age gap metric captured comparable prognostic information to its constituent biomarkers, we performed a complete case analysis restricted to participants with all four CT biomarkers available. We compared discrimination and calibration for non-lung cancer death across three Fine-Gray competing risks models: (1) clinical variables only (as above); (2) clinical variables plus age gap; and (3) clinical variables plus the four raw CT biomarkers as separate continuous parameters.

### Statistical methods and software

Results are presented according to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis: Artificial Intelligence (TRIPOD+AI) reporting guidelines<sup>21</sup>. Statistical analyses were conducted using R version 4 (R 4.4.2). R packages used included 'tidyverse', 'gt', 'gtsummary', 'mice', 'pmsampsize', 'tidycmprsk', 'riskRegression', 'survminer', and 'survival.' Further details on methods of regression of clinical features with aging metrics, feature selection, survival modelling, sample size calculation, and testing differences in concordance indices are provided in Supplementary Materials.

## Results

### Clinical characteristics

The final analysis included 12,478 participants, of whom 5,312 (43%) were female (Table 1). The median age at study enrollment was 65 years (interquartile range, IQR: 60–70). A total of 29,745 scans were analyzed across all time points. During a median follow-up of 5.0 years (IQR: 4.7–5.2) representing 58,284 person-years at risk, 593 participants (4.8%) were diagnosed with lung cancer and 966 (7.9%) died from any cause. Of the deaths, 851 (88%) were recorded as non-lung cancer-related deaths and 115 (12%) as lung cancer-related deaths.

### Disease course mapping

Three imaging progression models were fitted in an IECV scheme using data from all available timepoints (Figure 1). Individual estimates of age gap were inferred on baseline data from participants in the held-out region. The final form follows the cumulative distribution function after which the biomarkers were normalized, representing the average progression where the population mean of each biomarker is centered around 0.5 (Figure 2). All models demonstrated acceptable residual noise after fitting and consistency across folds (Supplementary Materials). The range of predicted age gap was –21.6 to 21.3 years (IQR –5 to +5 years).

### Clinical and demographic associations of age gap

The relationship between age gap and chronological age is shown in Figure 3. Participants who died during follow-up had a positive age gap compared with those who remained alive (mean: +2.2 years [standard deviation, SD 6.9] vs -0.2 years [SD 6.9], 95% CI 1.8-2.8,  $p < 0.001$ ). This difference was driven primarily by non-lung cancer-related deaths rather than lung cancer-related deaths. Participants who died from non-lung cancer causes had a mean age gap of +2.4 years (SD 6.9), higher than survivors (difference: 2.5 years, 95% CI: 2.1–3.0,  $p < 0.001$ ). In contrast, participants who had lung cancer related deaths had a mean age gap of +0.6 years (SD 7.1), which did not differ from survivors (difference: 0.6 years, 95% CI: -0.7 to

2.0,  $p=0.3$ ). Multivariable linear regressions examining associations between clinical characteristics and age gap are described in Supplementary Materials (Supplementary Figure 1, Supplementary Table 6).

## Association of age gap with competing outcomes

### Cumulative incidence of lung cancer and non-lung cancer death

To help visualise the cumulative incidence for both lung cancer diagnosis and non-lung cancer death (Figure 4), participants were stratified into quartiles based on age gap: 1st ( $<-5.1$  years), 2nd ( $-5.1-0$  years), 3rd ( $0-5.1$  years) and 4th ( $>5.1$  years). Age gap for the whole cohort ranged from  $-21.6$  to  $21.3$  years.

The 5-year cumulative incidence of lung cancer diagnosis was similar across age gap quartiles, ranging from 4.5% (95% CI: 3.8–5.3%) in quartile 1 to 5.7% in quartile 4 (95% CI: 4.9–6.6%,  $p=0.09$ , Gray's test). In contrast, the 5-year cumulative incidence of non-lung cancer death ranged from 3.8% (95% CI: 3.2–4.5%) in quartile 1, to 10% in quartile 4 (95% CI: 8.9–11%,  $p<0.001$ , Gray's test). (Supplementary Table 2)

### Subdistribution hazard models

Multivariable Fine-Gray competing risk models adjusted for clinical variables and age gap (Table 2) demonstrated that age gap was strongly associated with non-lung cancer death (sHR per 10 years: 2.0, 95% CI: 1.8–2.3,  $p<0.001$ ), but this effect was attenuated with lung cancer diagnosis (sHR: 1.3, 95% CI: 1.2–1.5,  $p<0.001$ ). For comparison, each 10-year increase in chronological age showed a sHR of 2.8 (95% CI: 2.5–3.2,  $p<0.001$ ) for non-lung cancer death and 2.2 (95% CI: 1.9–2.5,  $p<0.001$ ) for lung cancer diagnosis in the same multivariable models. There was no interaction between chronological age and age gap for both non-lung cancer death ( $p=0.2$ ) and lung cancer diagnosis ( $p=0.7$ ). There were no significant sex-smoking status, sex-smoking amount, or sex-time since quitting smoking interactions for either outcome.

Illustrative risks for example participant profiles were derived, showing parameter combinations where a younger participant may have a risk profile with higher risk of non-cancer related death because of older biological age and a positive age gap

(Profile A, Figure 5), relative to an older participant with younger biological age and a negative age gap with a high risk of lung cancer (Profile B).

### Model discrimination and calibration

Model performance was evaluated at a 5-year time horizon for each cluster and pooled using a random-effects meta-analysis (Figure 1). The baseline model (including clinical variables only) achieved a pooled C-index of 0.69 (95% CI: 0.66–0.72) for non-lung cancer-related death across the three validation folds (Figure 6). Addition of age gap to the baseline model improved discrimination to a pooled C-index of 0.73 (95% CI: 0.70–0.75). The C-index difference was +0.04 (95% CI 0.03–0.04,  $p < 0.001$ ). For lung cancer diagnosis, the baseline model achieved a pooled C-index of 0.66 (95% CI: 0.63–0.68). With the addition of age gap, the pooled C-index was 0.67 (95% CI: 0.64–0.70, Supplementary Figure 2). Similar differences in concordance were seen at a shorter 2-year time horizon (Supplementary Figure 3 and 4).

Calibration plots comparing predicted versus observed cumulative incidence of non-lung cancer death at 5 years showed good agreement across the range of predicted risks (Supplementary Figure 7). The pooled calibration slope was 0.90 (95% CI 0.78–1.02) and intercept was -0.03 (95% CI -0.12–0.05, Supplementary Table 8). Calibration plots for lung cancer diagnosis at 5 years showed underestimation of risks at higher risk categories, although at lower ranges calibration appears acceptable (Supplementary Figure 8). Calibration slope and intercept for each IECV fold and pooled estimates for 5 years are presented in Supplementary Tables 7 and 8, and for 2 years in Supplementary Tables 9 and 10.

An additional subgroup analysis of discrimination and calibration was performed stratified by socioeconomic deprivation and ethnicity, using out-of-fold prediction from IECV, showing that discrimination did not differ by ethnicity or by deprivation. Calibration showed that there was underprediction of non-lung cancer related death for more deprived participants and non-White participants (Supplementary Materials).

### Comparison with individual imaging features and established tools

To test the adequacy of the one-dimensional summary of the raw biomarkers, model performance was compared in a complete case analysis ( $n=10,782$ , 86% of the full cohort) using the same IECV framework across three specifications: clinical variables only, clinical variables with age gap, and clinical variables with the four raw CT biomarkers. The pooled C-index for predicting non-lung cancer death was 0.70 (95% CI: 0.67–0.72) for the clinical-only model, 0.73 (95% CI: 0.71–0.76) for the model including age gap, and 0.73 (95% CI: 0.70–0.76) for the model with raw biomarkers (Supplementary Figure 5). Discrimination for lung cancer diagnosis is presented in Supplementary Figure 6. Pooled estimates of calibration slope and intercept, as well as calibration plots across IECV clusters, are reported in Supplementary Tables 11 and 12, and Supplementary Figures 9 and 10.

We also compared the performance of the model with age gap with other established risk prediction methods using comorbidity indices, the PLCO-comorbidity index (PLCO-ci) and the Life-Years Gained from Screening CT (LYFS-CT) models. We find that both PLCO-ci and LYFS-CT performed comparably to the model with clinical variables only in our dataset, and the model with age gap outperformed both PLCO-ci and LYFS-CT for the prediction of non-lung cancer related death (C-index difference +0.039 and +0.029 respectively, both  $p<0.001$ ).

### Decision curve analysis

Decision curve analysis, using out-of-fold predicted risks from the internal-external cross-validation, showed that the addition of age gap to a model with clinical variables had higher net benefit than the clinical-only model for non-lung cancer death across threshold probabilities of 1-20% at 5 years (Supplementary Figure 13); both models showed net benefit over a 'treat all' strategy above a 2% threshold and over 'treat none' up to a 19% threshold. The advantage of adding age gap was smaller and less consistent at 2 years, consistent with the smaller 2-year discrimination gain reported above (Supplementary Figure 14).

## Discussion

Here we explore the use of a disease progression modelling approach to provide an estimate of biological age using imaging biomarkers in lung cancer screening. The difference between an individual's biological and chronological age ('age gap') inferred from this disease course map model improves the prediction of non-cancer related death in competing risks models compared with models using clinical variables alone. Such predictions could inform individualised discussions about the balance of benefits and harms from continued screening, particularly for individuals with high competing mortality risk.

We show that combining multiple imaging biomarkers into one single metric using a disease progression model does not reduce the predictive power of the competing risk models compared to using the individual metrics themselves. This dimensionality reduction has multiple practical implications. First, it may have utility in prediction models or other applications where parameter numbers are a concern, for example in risk prediction models fit in small sample sizes. Second, the framing of the single metric of 'age gap', where imaging biomarkers from an individual place them relative to their peers of the same chronological age, has potential advantages for clinicians, researchers, participants, and patients<sup>22</sup>. Communicating scan findings from multiple systems to lung cancer screening participants concisely could mean better shared decision making when it comes to decisions around continuing screening and lifestyle interventions. Indeed, the concept of a "healthy age" will already be familiar to those undertaking routine health checks in the NHS, where estimates of a "healthy heart age" as part of cardiovascular risk assessment are provided<sup>23,24</sup>.

Frailty and comorbidity are common within cancer populations<sup>25</sup>, but also in screening populations<sup>3</sup>, where the complex interplay between frailty, screening uptake, and oncogenesis needs careful consideration<sup>26</sup>. In the United States, risk prediction using clinical data, such as the Prostate, Lung, Colorectal, and Ovarian comorbidity index (PLCO-ci), developed in the PLCO trial and validated in NLST, found that participants of the intermediate frailty benefitted the most from lung cancer screening<sup>4</sup>. Also in the US, the life-years gained from screening-computed tomography (LYFS-CT)<sup>5</sup>, a model for patient selection similarly developed using

clinical and comorbidity data, increased the total life expectancy gain, whilst preventing slightly fewer lung cancer deaths. LYFS-CT is recommended by the American College of Chest Physicians, although further studies highlighted the need for careful consideration of race and ethnicity in these models to avoid exacerbating inequalities in lung cancer screening<sup>27</sup>. We find that our model with age gap performs favourably compared to both PLCO-ci and LYFS-CT, improving upon discrimination in our dataset. Furthermore, we did not find evidence that our model showed any differences in discrimination by either ethnicity or deprivation, and its underprediction of non-lung cancer related death in participants not of White ethnicity and of higher socioeconomic deprivation would tend to reduce the chances of these participants being excluded from screening based on their predicted risks.

Previously published risk models incorporating scan-based markers into predicting risk of competing death have used logistic regression, which introduces bias when handling competing risks, as well as only being applicable at a fixed time horizon<sup>7</sup>. Our approach presents the continuous predictor of age gap, which improves risk prediction in a time-to-event model of competing risks, at multiple time horizons, overcoming these shortcomings. This more flexible approach allows for broader adaptation to different populations and use cases, although these will require recalibration and revalidation.

Importantly, 'biological age' as used here is not a direct physiological measurement, but a latent parameter (the individual age gap estimated by the disease course model) expressed in years. Its meaning is therefore inherently relative to the model and the population it was fitted on. Therefore, our biological age gap metric is normalised to the LCS population of older adults with significant smoking history, rather than the general population. This means that an age gap calculated for an LCS participant cannot meaningfully compare them to an age matched individual from the general population. An individual who appears biologically older than their screening cohort peers carries higher competing mortality risk within that screening context, regardless of how the entire cohort compares to the general population. This approach of risk stratification within lung cancer screening populations aligns with the clinical decision-making context where the metric will be applied.

Non-lung cancer related causes of death in LCS for participants are predominantly cardiovascular disease, COPD and other cancers<sup>1,2</sup>. Evidence for incidental findings in LCS such as vessel calcification remains limited to observational associations that these imaging biomarkers are predictors of cardiovascular events and mortality<sup>28</sup>. No interventional trials have demonstrated that detecting and acting on these findings during lung cancer screening improves outcomes. The UK National Screening Committee review<sup>29</sup> classified incidental findings as a harm of screening, noting substantial heterogeneity in definitions and lack of evidence for benefit. Until such evidence exists, the contribution of incidental findings to screening benefit-harm balance remains uncertain<sup>30</sup>. The approach highlighted here therefore is to improve targeted lung cancer screening itself. Even though the disease progression model uses all longitudinal data, all model estimates were derived from baseline scan data. This shows how imaging biomarkers could enable more personalised, risk-stratified decisions with data available after the first LCS LDCT scan regarding screening continuation or frequency which account for an individual's balance of potential benefit from lung cancer detection versus competing mortality risks. We provide an illustrative example, with two profiles using variables from real participants. A participant who is chronologically younger but biologically older (a positive age gap, Figure 5, Profile A) may have a competing mortality risk that is disproportionately elevated relative to their lung cancer risk, which could support a conversation about whether continued screening at the standard interval remains a shared priority, as opposed to an extended screening interval or broader comorbidity assessment outside the screening pathway. Conversely, a participant who is chronologically older but biologically younger (a negative age gap, Profile B) may derive greater relative benefit from continued screening than their chronological age alone would suggest. Presented alongside conventional lung cancer risk estimates at a routine screening appointments, these metrics could support a more individualised conversation about continuing, extending, or exiting the screening pathway than either risk estimate provides in isolation.

The feasibility of implementing our models is another potential advantage. Unlike end-to-end deep learning prediction models which require significant computational resource, the most computationally intense part of this workflow is segmentation. Commercially available computer assisted detection software already in use in LCS

programmes<sup>31,32</sup> perform similar segmentations, and the further application of these measurements for inference in a disease course map would require no more resource than a modern personal computer. Practically, this could be integrated into existing lung cancer screening workflows by running inference on the segmentation outputs already generated by commercial computer-assisted detection software at the point of the baseline scan, without requiring additional imaging or a separate visit. Future work could also explore whether incorporating additional CT-derived biomarkers could improve discrimination further within this disease course mapping framework.

There are important limitations to this study. First, the wider geographical transportability to other lung cancer screening settings is not explored here. Whilst the internal-external cross validation approach taken here uses the clustering inherent in our large lung cancer screening study to simulate natural geographical variation and come to more realistic estimates of model performance, this is still limited to three clusters in a relatively limited geographical area. Furthermore, all data were acquired within a single trial with standardised scanner models and imaging protocols, and model performance with different scanners, acquisition protocols, or in other screening populations remains untested and requires independent external validation before clinical use. Second, the negative association between mean age gap and chronological age suggests an intuitive explanation of a selection effect. However, showing this would require age-matched imaging data from the general population, which were not available. Third, it is possible that risk prediction models using more sophisticated image analysis techniques such as end-to-end deep learning could achieve better absolute performance. However, these models often reach barriers for clinical translation because of complex technical implementation or clinician and patient hesitancy due to interpretability and explainability of its outputs<sup>33</sup>. Finally, modelling the competing risk of non-lung cancer death as lung cancer diagnosis may overstate screening benefit to the extent that some diagnoses represent overdiagnosis. However, lung cancer deaths were substantially fewer than diagnoses in our cohort, and lacked power for precise estimates of performance when using lung cancer death as an alternative outcome.

## Conclusions

Imaging at baseline lung cancer screening can be used to construct a single biological age gap metric that improves competing risk prediction beyond chronological age and clinical variables. Quantifying individualised risk of non-lung cancer death could inform shared decision-making about screening continuation and enable risk-stratified screening strategies that better balance benefits and harms for individual participants.

## Tables

| Characteristic                                          | N = 12,478   |
|---------------------------------------------------------|--------------|
| Age at baseline, years                                  | 65 (60, 70)  |
| Sex, female                                             | 5,312 (43%)  |
| Ethnicity                                               |              |
| Asian, Asian British or Asian Welsh                     | 916 (7.3%)   |
| Black, Black British, Black Welsh, Caribbean or African | 565 (4.5%)   |
| Mixed or Multiple ethnic groups                         | 270 (2.2%)   |
| Other ethnic group                                      | 349 (2.8%)   |
| White                                                   | 10,378 (83%) |
| Current smoker (any tobacco)                            | 6,089 (49%)  |
| Total smoking duration, years                           | 44 (40, 48)  |
| Unknown                                                 | 3            |
| Weighted smoking amount, pack-years                     | 20 (15, 22)  |
| Previous diagnosis of COPD, bronchitis or emphysema     | 3,661 (29%)  |
| Maximum educational attainment                          |              |
| Completed a further degree (Masters, PhD or equivalent) | 622 (5.0%)   |
| Completed A Levels or equivalent                        | 1,344 (11%)  |
| Completed bachelor's degree or equivalent               | 1,539 (12%)  |
| Completed further education but not degree              | 1,057 (8.5%) |
| Completed GCSEs, O-Levels or equivalent                 | 2,961 (24%)  |
| Finished school at or before age 15                     | 4,955 (40%)  |
| Body mass index category, kg/m <sup>2</sup>             |              |
| Normal weight (18.5–24.9)                               | 3,631 (29%)  |
| Underweight ( $\leq 18.5$ )                             | 180 (1.4%)   |
| Overweight (25–29.9)                                    | 4,835 (39%)  |
| Obese ( $>30$ )                                         | 3,832 (31%)  |
| IMD rank quintile                                       |              |
| 1 (most deprived)                                       | 3,988 (32%)  |
| 2                                                       | 3,691 (30%)  |
| 3                                                       | 2,192 (18%)  |
| 4                                                       | 1,848 (15%)  |
| 5 (least deprived)                                      | 627 (5.1%)   |
| Unknown                                                 | 132          |
| Lung cancer diagnosed during screening                  | 593 (4.8%)   |
| Death at follow up                                      | 966 (7.9%)   |
| Unknown                                                 | 178          |

**Table 1. Characteristics of included participants.** Values are expressed as median (interquartile range) or number (%). COPD: Chronic obstructive pulmonary disease; IMD: Indices of Multiple Deprivation.

| Characteristic                          | LC diagnosis |            |         | Non LC-related death |            |         |
|-----------------------------------------|--------------|------------|---------|----------------------|------------|---------|
|                                         | sHR          | 95% CI     | p-value | sHR                  | 95% CI     | p-value |
| Age at baseline, 10y                    | 2.15         | 1.86, 2.47 | <0.001  | 2.81                 | 2.45, 3.21 | <0.001  |
| Age gap , 10y                           | 1.32         | 1.16, 1.51 | <0.001  | 2.02                 | 1.79, 2.27 | <0.001  |
| Sex, female                             | 1.08         | 0.92, 1.28 | 0.4     | 0.58                 | 0.49, 0.67 | <0.001  |
| Body mass index, kg/m <sup>2</sup>      | 0.96         | 0.94, 0.97 | <0.001  | 0.99                 | 0.97, 1.01 | 0.3     |
| Maximum educational attainment          | 0.93         | 0.89, 0.99 | 0.013   | 0.89                 | 0.85, 0.93 | <0.001  |
| COPD, bronchitis or emphysema diagnosis | 1.24         | 1.05, 1.47 | 0.013   | 1.47                 | 1.27, 1.70 | <0.001  |
| Family history of lung cancer           | 1.71         | 1.41, 2.07 | <0.001  | 0.87                 | 0.70, 1.07 | 0.2     |
| Any prior cancer                        | 1.51         | 1.24, 1.84 | <0.001  | 1.22                 | 1.01, 1.47 | 0.037   |
| Current smoker (any tobacco)            | 1.31         | 1.06, 1.62 | 0.014   | 1.18                 | 0.97, 1.42 | 0.094   |
| Years since stopping all tobacco        | 1.00         | 0.98, 1.01 | 0.8     | 0.99                 | 0.98, 1.01 | 0.3     |
| Weighted smoking amount, pack-years     | 1.01         | 1.00, 1.01 | 0.069   | 1.01                 | 1.00, 1.01 | 0.046   |

**Table 2. Fine-Gray competing risks models for lung cancer diagnosis and non-lung cancer-related death.** LC: lung cancer; sHR: subdistribution hazard ratio; CI: confidence interval; COPD: Chronic obstructive pulmonary disease.

## Figures

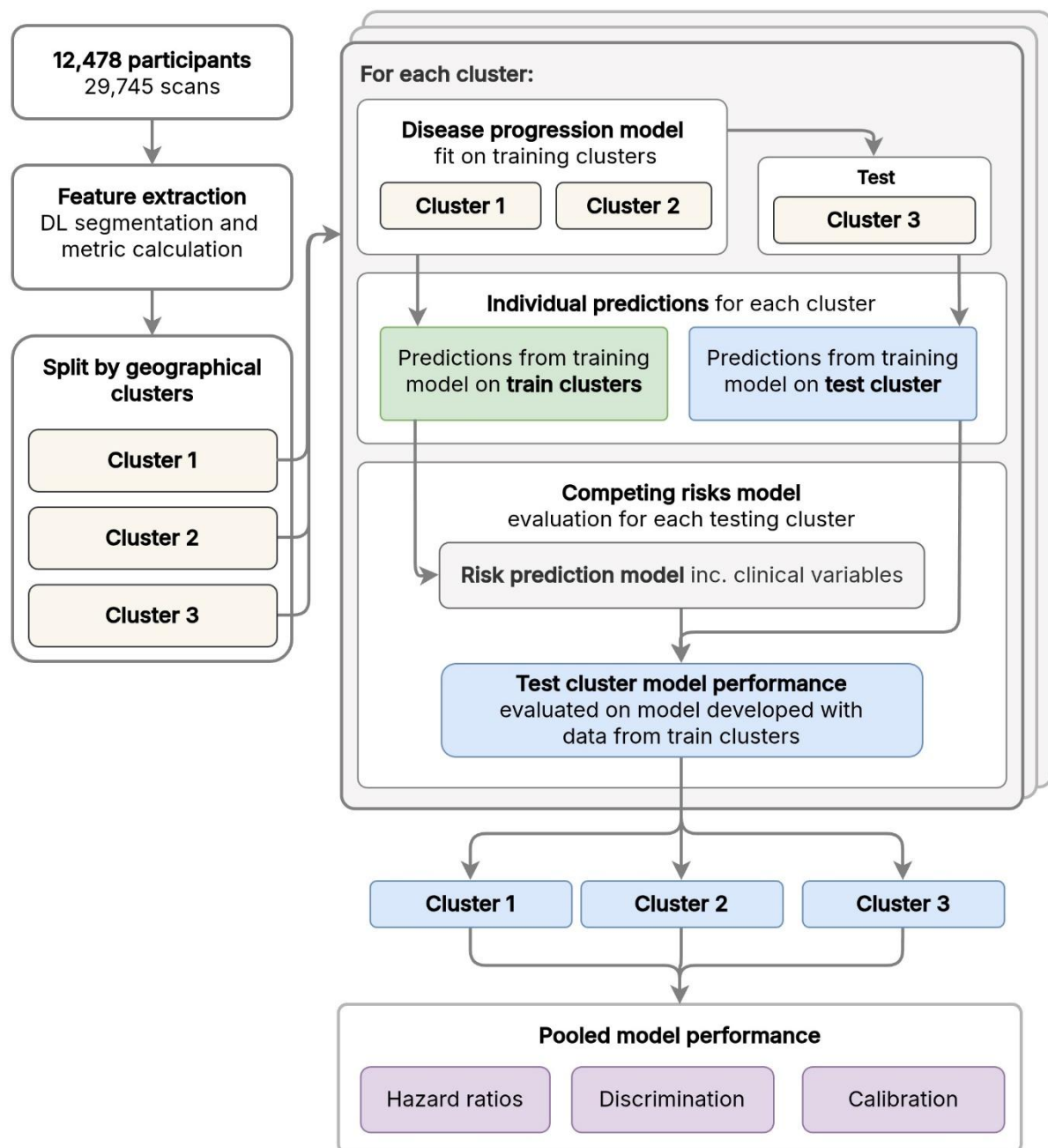


**Figure 1. Summary of model development and inference process for individual predictions.** Disease progression and risk prediction models are fitted within each geographical cluster, following an internal-external cross validation framework. DL: Deep Learning.

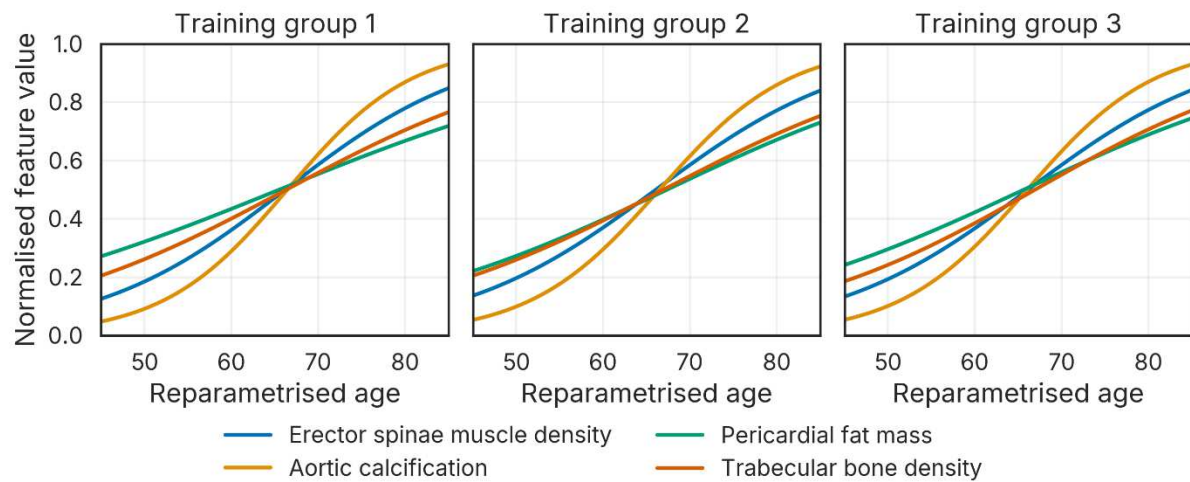


**Figure 2. Average trajectory of individual biomarker evolution by reparameterised age as estimated by disease course map model, by IECV training groups.** Training group 1 is trained on IECV clusters 1 and 2; training group 2 to clusters 1 and 3; and training group 3 to clusters 2 and 3. Features are normalised between 0 and 1 using an empirical cumulative distribution function, and signs appropriately changed so biomarkers increase with age. IECV: Internal-external cross validation.

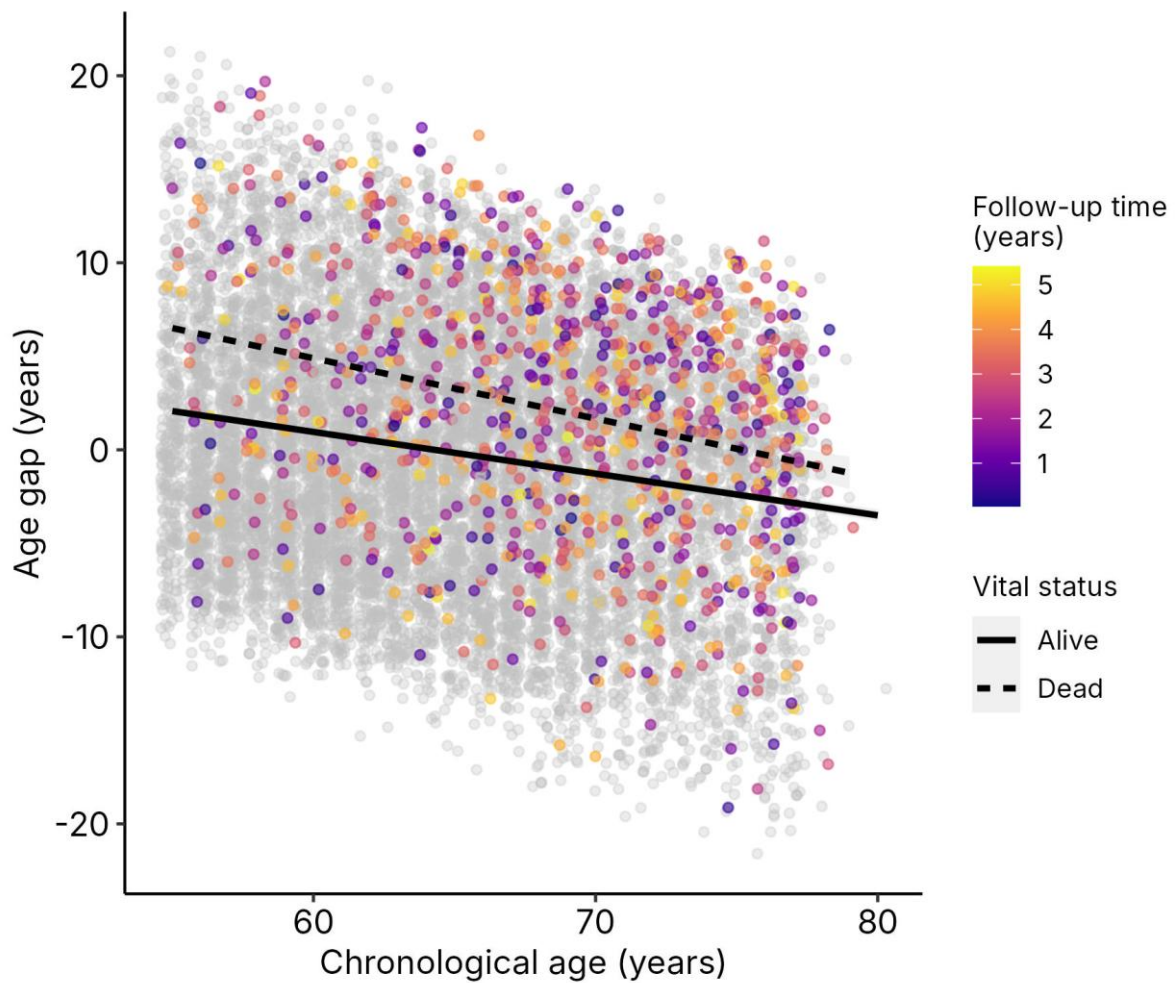


**Figure 3. Age gap versus chronological age.** Each point represents one participant. For participants who died, follow up time is represented in a color scale. Best fit line fit through participants who died at follow up (dashed line) and participants alive at follow up (solid line).

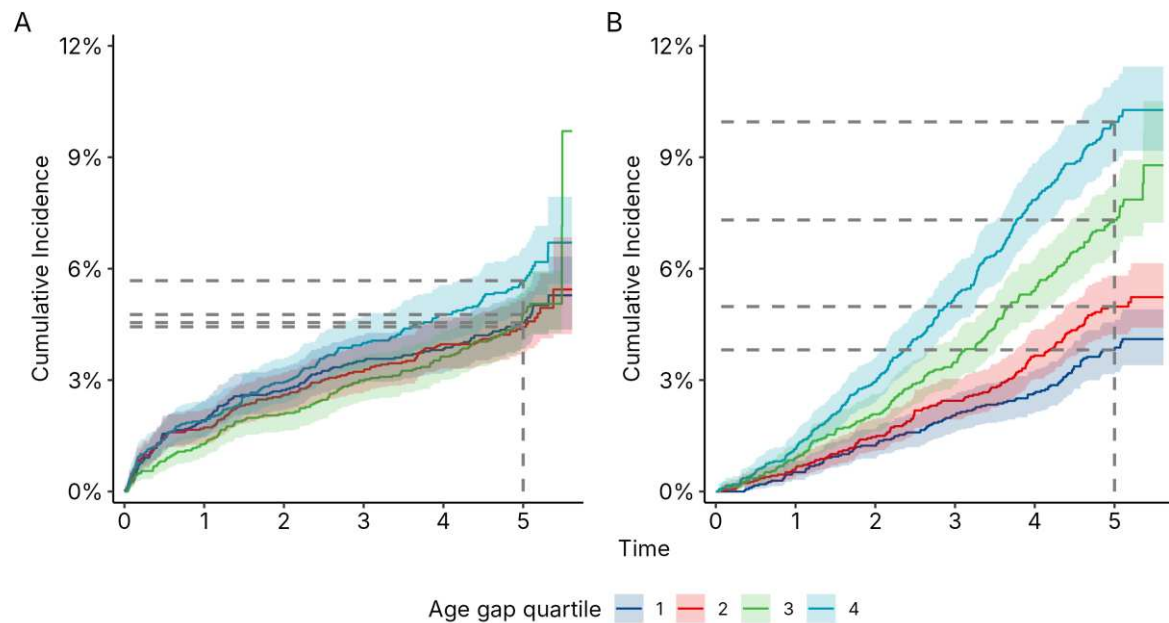


**Figure 4. Cumulative incidence of lung cancer diagnosis (A) and non-lung cancer related death (B) stratified by age gap.** Age gap is stratified into quartiles, where 1 is the lowest age gap and 4 is the highest. Dashed lines indicate cumulative incidence at 5 years.

## Biological age and competing risk of death in lung cancer screening

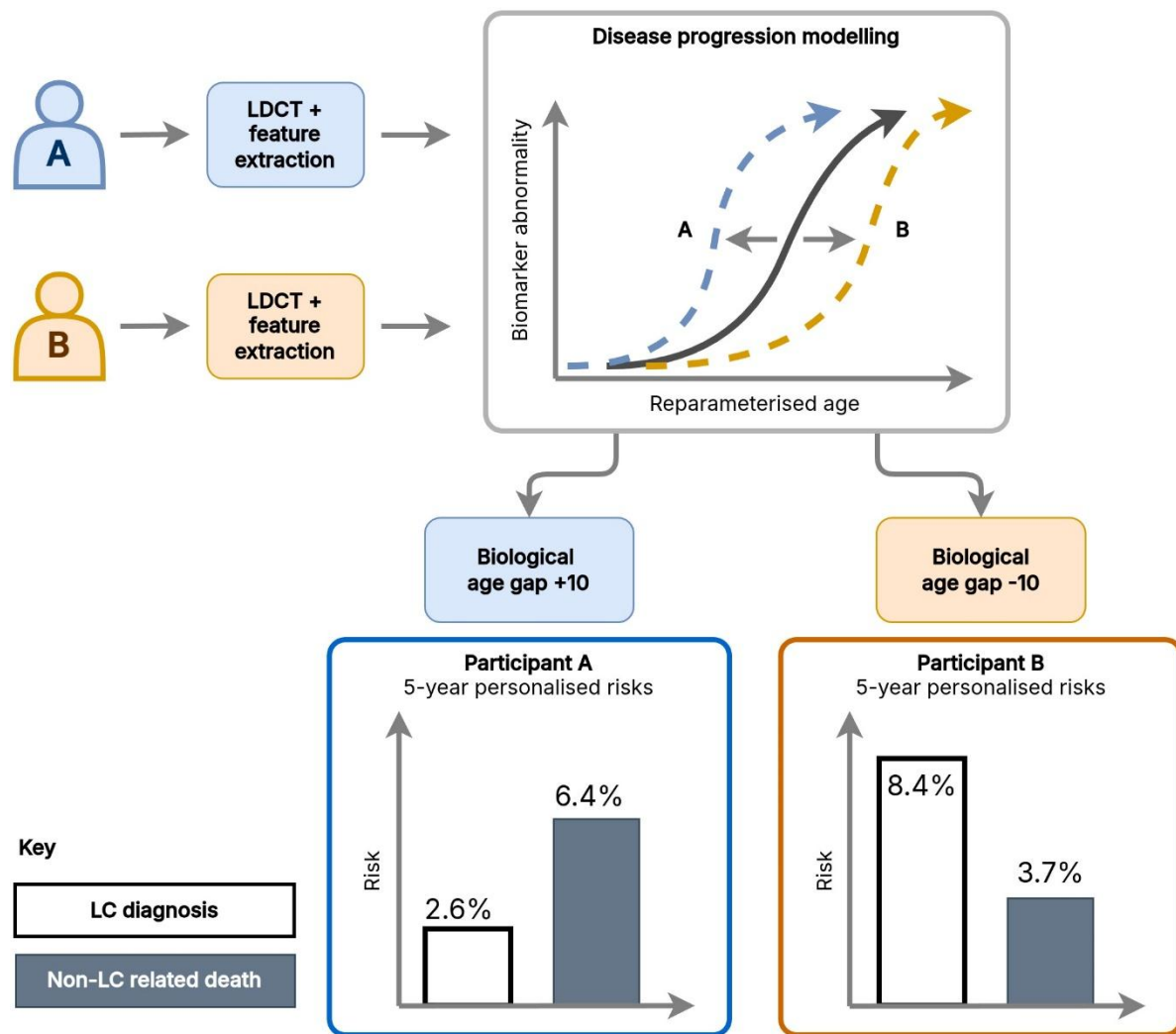


**Figure 5. Illustrative examples of individual predictions of competing risks at 5 years.** Both risks from Fine-Gray models. Participant A is a 55-year-old male with no family history of lung cancer, and age gap +10 years. Participant B is a 70-year-old female with a family history of lung cancer, and age gap -10 years. Both participants have a baseline body mass index of 27 kg/m<sup>2</sup>, a diagnosis of chronic obstructive pulmonary disease, no personal history of cancer, are former smokers, quit smoking 5 years ago, 30 pack-year smoking history, and left school before age 15. LDCT: Low dose computed tomography; LC: Lung cancer.

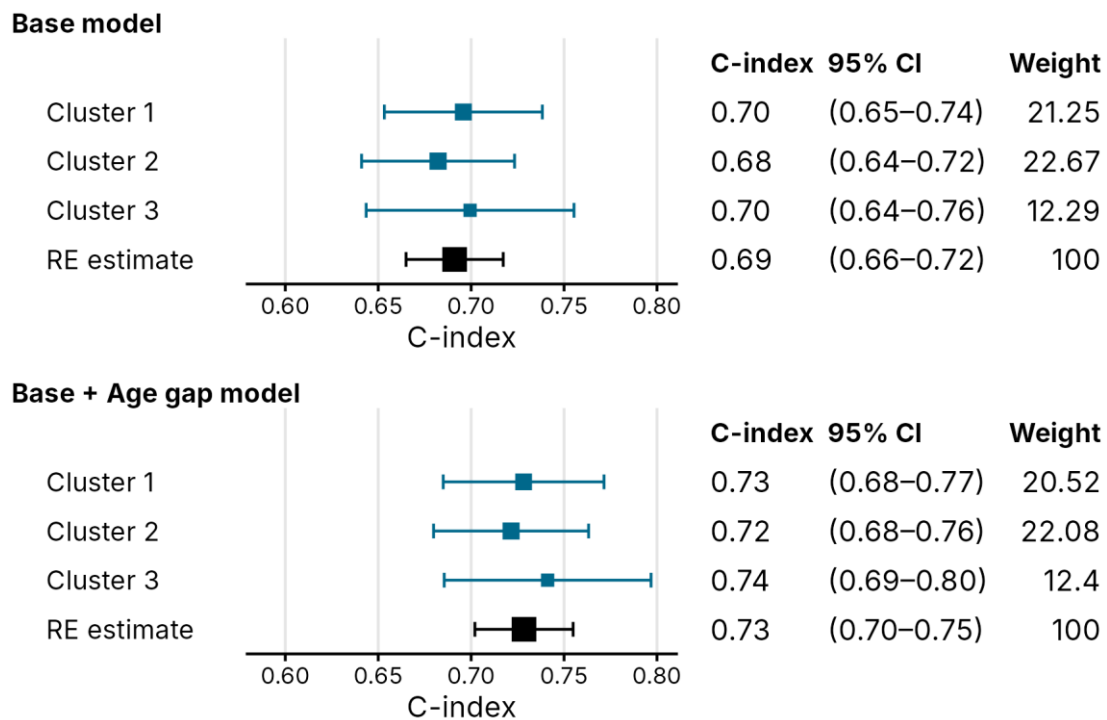


**Figure 6. Forest plot of competing risks model concordance for non-lung cancer death at 5-year time horizon by IECV clusters and random effects meta-analyzed estimates.** C-index: Concordance index; CI: confidence interval.

## References

1. Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening. *New England Journal of Medicine* **365**, 395–409 (2011).
2. de Koning, H. J. *et al.* Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. *New England Journal of Medicine* **382**, 503–513 (2020).
3. Almatrafi, A. *et al.* Impact of frailty and comorbidity on initial response to lung cancer screening invitation and low-dose CT screening uptake: Findings from the Yorkshire Lung Screening Trial. *J. Med. Screen.* <https://doi.org/10.1177/09691413251315087>;SUBPAGE:STRING:FULL (2025) doi:10.1177/09691413251315087;SUBPAGE:STRING:FULL.
4. Gendarme, S. *et al.* Impact of Comorbidities on the Mortality Benefits of Lung Cancer Screening: A Post-Hoc Analysis of the PLCO and NLST Trials. *Journal of Thoracic Oncology* **20**, 565–576 (2025).
5. Cheung, L. C., Berg, C. D., Castle, P. E., Katki, H. A. & Chaturvedi, A. K. Life-Gained–Based Versus Risk-Based Selection of Smokers for Lung Cancer Screening. <https://doi.org/10.7326/M19-1263> **171**, 623–632 (2019).
6. Mikhael, P. G. *et al.* Sybil: A Validated Deep Learning Model to Predict Future Lung Cancer Risk from a Single Low-Dose Chest Computed Tomography. *Journal of Clinical Oncology* **41**, 2191–2200 (2023).
7. Schreuder, A. *et al.* Scan-based competing death risk model for re-evaluating lung cancer computed tomography screening eligibility. *European Respiratory Journal* **59**, (2022).
8. Wasserthal, J. *et al.* TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images. <https://doi.org/10.1148/ryai.230024> **5**, (2023).

9. Engelke, K. *et al.* The effect of ageing on fat infiltration of thigh and paraspinal muscles in men. *Aging Clinical and Experimental Research* 2022 34:9 **34**, 2089–2098 (2022).
10. Chen, H., Zhou, X., Fujita, H., Onozuka, M. & Kubo, K. Y. Age-Related Changes in Trabecular and Cortical Bone Microstructure. *Int. J. Endocrinol.* **2013**, 213234 (2013).
11. Pescatore, L. A., Gamarra, L. F. & Liberman, M. Multifaceted Mechanisms of Vascular Calcification in Aging. *Arterioscler. Thromb. Vasc. Biol.* **39**, 1307–1316 (2019).
12. Ni, X. *et al.* Correlation between the distribution of abdominal, pericardial and subcutaneous fat and muscle and age and gender in a middle-aged and elderly population. *Diabetes, Metabolic Syndrome and Obesity* **14**, 2201–2208 (2021).
13. Debray, T. P. A., Moons, K. G. M., Ahmed, I., Koffijberg, H. & Riley, R. D. A framework for developing, implementing, and evaluating clinical prediction models in an individual participant data meta-analysis. *Stat. Med.* **32**, 3158–3180 (2013).
14. Takada, T. *et al.* Internal-external cross-validation helped to evaluate the generalizability of prediction models in large clustered datasets. *J. Clin. Epidemiol.* **137**, 83–91 (2021).
15. Steyerberg, E. W. & Harrell, F. E. Prediction models need appropriate internal, internal-external, and external validation. *J. Clin. Epidemiol.* **69**, 245 (2015).
16. Young, A. L. *et al.* Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference. *Nature Communications* 2018 9:1 **9**, 4273- (2018).
17. Young, A. L. *et al.* Disease Progression Modeling in Chronic Obstructive Pulmonary Disease. <https://doi.org/10.1164/rccm.201908-1600OC> **201**, 294–302 (2020).

18. Koval, I. *et al.* AD Course Map charts Alzheimer's disease progression. *Sci. Rep.* **11**, 1–16 (2021).
19. Van Geloven, N. *et al.* Validation of prediction models in the presence of competing risks: a guide through modern methods. *BMJ* **377**, (2022).
20. Putter, H., Fiocco, M. & Gekus, R. B. Tutorial in biostatistics: Competing risk and multi-state models. *Stat. Med.* **26**, 2389–2430 (2007).
21. Collins, G. S. *et al.* TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* **385**, (2024).
22. Haugg, F. *et al.* Imaging biomarkers of ageing: a review of artificial intelligence-based approaches for age estimation. *Lancet Healthy Longev.* **6**, 100728 (2025).
23. Hippisley-Cox, J., Coupland, C. & Brindle, P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ* **357**, (2017).
24. Calculate your heart age - NHS. <https://www.nhs.uk/health-assessment-tools/calculate-your-heart-age>.
25. Komici, K. *et al.* Frailty in Patients With Lung Cancer: A Systematic Review and Meta-Analysis. *Chest* **162**, 485–497 (2022).
26. Perrotta, F. *et al.* Immune Checkpoint Blockade for Advanced NSCLC: A New Landscape for Elderly Patients. *Int. J. Mol. Sci.* **20**, 2258 (2019).
27. Landy, R. *et al.* Methods for Using Race and Ethnicity in Prediction Models for Lung Cancer Screening Eligibility. *JAMA Netw. Open* **6**, e2331155–e2331155 (2023).
28. Chiles, C. *et al.* Association of Coronary Artery Calcification and Mortality in the National Lung Screening Trial: A Comparison of Three Scoring Methods. <https://doi.org/10.1148/radiol.15142062> **276**, 82–90 (2015).

29. Lung cancer - UK National Screening Committee (UK NSC) - GOV.UK.  
<https://view-health-screening-recommendations.service.gov.uk/lung-cancer/>.
30. Henderson, L. M. *et al.* Lung Cancer Screening and Incidental Findings: A Research Agenda: An Official American Thoracic Society Research Statement. <https://doi.org/10.1164/rccm.202501-0011ST> **211**, 436–451 (2025).
31. Veolity - a brand of MeVis Medical Solutions AG: Home.  
<https://www.veolity.com/>.
32. Murchison, J. T. *et al.* Validation of a deep learning computer aided system for CT based lung nodule detection, classification, and growth rate estimation in a routine clinical population. *PLoS One* **17**, e0266799 (2022).
33. Imrie, F., Davis, R. & van der Schaar, M. Multiple stakeholders drive diverse interpretability requirements for machine learning in healthcare. *Nature Machine Intelligence* **2023 5:8 5**, 824–829 (2023).

## Acknowledgements

This research was funded in whole or in part by the Wellcome Trust (227835/Z/23/Z; 227341/Z/23/Z), the Rosetrees Trust (PGL25-Full/100046, CF-2023-M-2\105) and the Robert Luff Foundation (PGL25-Full/100046). DC is supported by a Clinical Research Training Fellowship from the Alliance for Cancer Early Detection (ACEDAS-2023/100003). SA is supported by the International Alliance for Cancer Early Detection, an alliance between Cancer Research UK [EDDAPA-2023/100002], Canary Center at Stanford University, the University of Cambridge, OHSU Knight Cancer Institute, University College London and the University of Manchester. JJ and SMJ were also supported by the NIHR UCLH Biomedical Research Centre. NPO is a UKRI Future Leaders Fellow (MR/S03546X/1, MR/X024288/1). This project has also been made possible in part by grant number CZIF2024-009938 from the Chan Zuckerberg Initiative Foundation. The SUMMIT Study is funded by GRAIL LLC through a research grant awarded to SMJ as Principal Investigator.

**SUMMIT consortium:** Aashna Samson<sup>4</sup>, Adefiola Olabode<sup>4</sup>, Alberto Villanueva<sup>4</sup>, Ali Mohammed<sup>13</sup>, Alice Cotton<sup>4</sup>, Andrew Crossingham<sup>4</sup>, Andrew Perugia<sup>22</sup>, Anjali Ketkar<sup>4</sup>, Anna Sikorska<sup>4</sup>, Anthony Edey<sup>4</sup>, Antonette Andrews<sup>4</sup>, April Neville<sup>4</sup>, Burcu Ozaltin<sup>20</sup>, Camilla Mitchell<sup>4</sup>, Catherine Nestor<sup>4</sup>, Celia Snell<sup>4</sup>, Charlie Sayer<sup>4</sup>, Charlotte Cash<sup>7</sup>, Chimtom Nwaosu<sup>4</sup>, Christine Hosein<sup>4</sup>, Claire Levermore<sup>4</sup>, Columbus Ife<sup>4</sup>, Derya Ovayolu<sup>4</sup>, Dominique Arancon<sup>4</sup>, Domminique Arancon<sup>4</sup>, Eleanor Hellier<sup>4</sup>, Elena Stefan<sup>4</sup>, Elodie Murali<sup>4</sup>, Esther Arthur-Darkwa<sup>4</sup>, Ethaar El-Emir<sup>4</sup>, Fabia Ruggiero<sup>4</sup>, Fahiza Chowdhury<sup>4</sup>, Fahmida Hoque<sup>4</sup>, Fanta Bojang<sup>4</sup>, Farhida Begum<sup>4</sup>, Geoff Bellingan<sup>4</sup>, Georgia Bullock<sup>4</sup>, Graham Robinson<sup>4</sup>, Gulistan Keskinbicak<sup>4</sup>, Gulten Geneci<sup>4</sup>, Harriet Sorrell<sup>4</sup>, Hasti Robbie<sup>4</sup>, Helen Kiconco<sup>4</sup>, Helen Victoria Bowyer<sup>4</sup>, Henna Makwana<sup>4</sup>, Hina Pervez<sup>4</sup>, Houda Benmebarek<sup>4</sup>, James Rusius<sup>22</sup>, Jane Rowlands<sup>4</sup>, Janine Zylstra<sup>4</sup>, Jeannie Eng<sup>4</sup>, Jhanara Begum<sup>4</sup>, Jon Teague<sup>2</sup>, Joseph Jacob<sup>20</sup>, Judy Airebamen<sup>4</sup>, Julian McKee<sup>4</sup>, Karen Parry-Billings<sup>4</sup>, Karen Sennett<sup>23</sup>, Kate Davies<sup>1</sup>, Kate Gowers<sup>1</sup>, Kaylene Phua<sup>4</sup>, Kim Sorley<sup>4</sup>, Kitty Chan<sup>2</sup>, Kylie Gyertson<sup>4</sup>, Laila Ahmed<sup>4</sup>, Laura Farrelly<sup>2</sup>, Laura Green<sup>4</sup>, Lucas Clarke<sup>4</sup>, Lynsey Gallagher<sup>4</sup>, Magali Taylor<sup>4</sup>, Malavika Suresh<sup>2</sup>, Mamta Ruparel<sup>1</sup>, Marie Calligaris<sup>4</sup>, Marilyn Ijeomah-Orji<sup>4</sup>, Mark Clark<sup>4</sup>, Maureen Browne<sup>4</sup>, Mehran Azimbagirad<sup>20</sup>, Miguel

Ferreira<sup>4</sup>, Milica Rajkov<sup>4</sup>, Moksud Miah<sup>4</sup>, Monika Sommerrey<sup>4</sup>, Navinah Nundlall<sup>4</sup>, Nazim Miah<sup>4</sup>, Neil Songsong<sup>4</sup>, Nicholas Beech<sup>4</sup>, Nick Screatton<sup>19</sup>, Nikita Sharma<sup>4</sup>, Olufunmilola Ajayi<sup>4</sup>, Oluwadara Akerewusi<sup>4</sup>, Patricia Castello<sup>4</sup>, Paul Hickey<sup>4</sup>, Paul Robinson<sup>4</sup>, Qin Neville<sup>4</sup>, Rabiya Patel<sup>4</sup>, Rachael Sarpong<sup>2</sup>, Ricardo McEwen<sup>4</sup>, Rosie Dapaah Tweneboah<sup>4</sup>, Rowan Faiers<sup>4</sup>, Sahana Uthayakumar<sup>4</sup>, Sam Hare<sup>4</sup>, Samanjit Hare<sup>7</sup>, Sara Lock<sup>8</sup>, Sebri Hasaballa<sup>4</sup>, Shaira Hassan<sup>4</sup>, Sheetal Karavadra<sup>4</sup>, Shrinkhala Dawadi<sup>2</sup>, Shummi Begum<sup>4</sup>, Simranjit Mehta<sup>4</sup>, Sofia Nnorom<sup>4</sup>, Stephen Ellis<sup>4</sup>, Stuart Marnham<sup>4</sup>, Sumiya Mahey<sup>4</sup>, Suraiya Sharmin<sup>4</sup>, Syeda Anam Hussain<sup>4</sup>, Sylvia Patricia Enes<sup>4</sup>, Tania Anastasiadis<sup>21</sup>, Tania Anastasiadis<sup>21</sup>, Tanita Limani<sup>4</sup>, Tanya Patrick<sup>1</sup>, Terry O'Shaughnessy<sup>13</sup>, Thea Buchan<sup>4</sup>, Theepan Visakan<sup>4</sup>, Thomas Callender<sup>1</sup>, Tracy Odigie<sup>4</sup>, Trisha Enes<sup>4</sup>, Tunku Aziz<sup>13</sup>, Urja Patel<sup>4</sup>, Vicky Bowyer<sup>4</sup>, Ya-Fei Li<sup>4</sup>, Yee Chin Lee<sup>4</sup>, Zaheer Mangera<sup>12</sup>, Zahra Awan<sup>4</sup>, Zahra Hanif<sup>4</sup>, Zelena Aziz<sup>4</sup>.

Affiliations: 1- Lungs For Living Research Centre, UCL Respiratory, University College London, London; 2- CRUK & UCL Cancer Trials Centre, University College London, London; 3- Centre for Cancer Screening, Prevention, Detection and Early Diagnosis, Wolfson Institute of Population Health, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London; 4- University College London Hospitals NHS Foundation Trust, London; 5- Royal Brompton and Harefield NHS Foundation Trust, London; 6- National Heart and Lung Institute, Imperial College, London; 7- Royal Free London NHS Foundation Trust, London; 8- Whittington Health NHS Trust, London; 9- Barking, Havering and Redbridge University Hospitals NHS Trust, Essex; 10- Homerton University Hospital Foundation Trust, London; 11- The Princess Alexandra Hospital NHS Trust, Essex; 12- North Middlesex University Hospital NHS Trust, London; 13- Barts Health NHS Trust, London; 14- North Bristol NHS Trust, Bristol; 15- Royal United Hospitals Bath NHS Foundation Trust, Bath; 16- Surrey and Sussex Healthcare NHS Trust, Surrey; 17- King's College Hospital NHS Foundation Trust, London; 18- University Hospitals Sussex NHS Foundation Trust, Sussex; 19- Royal Papworth Hospital NHS Foundation Trust, Cambridge; 20-Satsuma Lab, Hawkes Institute, London; 21- Tower Hamlets Clinical Commissioning Group, London; 22- Noclor Research Support, London; 23- Islington Clinical Commissioning Group, London.

## Declarations and Conflicts of Interest:

JJ declares consultancy fees from Boehringer Ingelheim, F. Hoffmann-La Roche, GlaxoSmithKline, NHSX; fees from advisory Boards for Boehringer Ingelheim, F. Hoffmann-La Roche; lecture fees from Boehringer Ingelheim, F. Hoffmann-La Roche, Takeda; grant funding from GlaxoSmithKline, Wellcome Trust, Microsoft Research, Gilead Sciences, Chan Zuckerberg Initiative and UK patent application numbers: 2113765.8 and GB2211487.0.

SMJ is supported by CRUK programme grant (EDDCPGM\100002), and MRC Programme grant (MR/W025051/1). SMJ receives support from the CRUK Lung Cancer Centre of Excellence (C11496/ A30025) and the CRUK City of London Centre, the Rosetrees Trust, the Roy Castle Lung Cancer foundation, the Longfonds BREATH Consortia, MRC UKRMP2 Consortia, the Garfield Weston Trust and University College London Hospitals Charitable Foundation. SMJ's work is supported by a Stand Up To Cancer-LUNGevity- American Lung Association Lung Cancer Interception Dream Team Translational Research Grant and Johnson and Johnson (grant number: SU2C-AACR-DT23-17 to S.M. Dubinett and A.E. Spira). Stand Up To Cancer is a division of the Entertainment Industry Foundation. Research grants are administered by the American Association for Cancer Research, the Scientific Partner of SU2C. This work was partly undertaken at UCLH/UCL who received a proportion of funding from the Department of Health's NIHR Biomedical Research Centre's funding scheme. SMJ has received fees for advisory board membership in the last three years from Bard1 Lifescience. He has received grant income from GRAIL Inc. He is an unpaid member of a GRAIL advisory board. He has received lecture fees for academic meetings from Cheisi and Astra Zeneca. His wife works for Astra Zeneca.

JRH has received grant support, and payment for educational and advisory work, personally and to his institution from AstraZeneca, Boehringer Ingelheim, Chiesi, GlaxoSmithKline, Sanofi, Sanofi-Regeneron, Takeda.

AN has received a consultation fee from MSD and honoraria speaking engagements from Mevis Medical Solutions. AN is an Honorary Executive Committee member for the British Society of Thoracic Imaging, Co-chair of the British Thoracic Society Pulmonary Nodule Guideline Development Group, Lung Taskforce member for the

Biological age and competing risk of death in lung cancer screening

British Lung Foundation and clinical lead for the NHS England Targeted Lung Health Checks Program.

ACPS received honoraria from Roche and educational fees from Amgen.

DA is Director and a shareholder of Queen Square Analytics.

**Competing interest declaration:** “All authors have completed the Unified Competing Interest form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available on request from the corresponding author) and declare that (1) no authors have support from a company for the submitted work; (2) SMJ, JJ, AN have specified relationships with one or more of: AstraZeneca, Boehringer Ingelheim, Chiesi, GlaxoSmithKline, F. Hoffmann-La Roche, Gilead Sciences, Sanofi, Sanofi-Regeneron, Microsoft Research, Johnson and Johnson, Bard1 Lifescience , GRAIL, Mevis Medical Solutions, MSD, Queen Square Analytics, Takeda that might have an interest in the submitted work in the previous 3 years; (3) SMJ's spouse works for AstraZeneca, but no other coauthors spouses, partners, or children have financial relationships that may be relevant to the submitted work; and (4) All authors have no non-financial interests that may be relevant to the submitted work.”

**Role of study sponsors:** The study sponsors did not have a role in design or in the collection, analysis, and interpretation of data.

**Other declarations:** The investigators were independent from the funders; JJ had full access to the data and can take responsibility for the integrity of the data and the accuracy of the data analysis; JJ (the manuscript's guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

**Data sharing:** No additional data available.

## Author Contribution

### CRedit (Contributor Roles Taxonomy) Statement

Daryl O. Cheng<sup>‡</sup> : Conceptualisation, methodology, software, validation, formal analysis, writing (original draft, review and editing), visualization; John McCabe: methodology, data curation, writing (review and editing), project administration; Tanya Patrick: writing (review and editing); Robert Chapman: writing (review and editing); Shanshan Wang: data curation; Tianqi Yang: data curation; Adam Szmul: data curation; Mehran Azimbagirad: data curation; Shahab Aslani: data curation; Monica L. Mullin: writing (review and editing); Sindhu Bhaarrati Naidu: writing (review and editing); Chuen Ryan Khaw: writing (review and editing); Alfred CP So: writing (review and editing); SUMMIT consortium\*: resources; Neal Navani: writing (review and editing); Arjun Nair: writing (review and editing); John R Hurst: writing (review and editing); Andre Altmann: writing (review and editing), conceptualization; Adam Pennycuik: writing (review and editing), conceptualization; Daniel C. Alexander: resources; Neil P. Oxtoby: methodology, writing (review and editing); Alexandra L. Young: methodology, writing (review and editing); Matthew Sperrin: conceptualisation, methodology, writing (review and editing), supervision; Sam M. Janes <sup>†</sup>: resources, writing (review and editing), funding acquisition, supervision; and Joseph Jacob <sup>†</sup>: resources, writing (review and editing), funding acquisition, supervision.

\*: listed in acknowledgements; <sup>†</sup>: represents joint senior authorship; <sup>‡</sup>:

Corresponding author: daryl.cheng@ucl.ac.uk

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SUMMITageingSupplementr2clean.docx](#)