



# Scan-based competing death risk model for re-evaluating lung cancer computed tomography screening eligibility

Anton Schreuder <sup>1</sup>, Colin Jacobs <sup>1</sup>, Nikolas Lessmann <sup>1</sup>, Mireille J.M. Broeders<sup>2,3</sup>, Mario Silva<sup>4,5</sup>, Ivana Išgum<sup>6,7</sup>, Pim A. de Jong<sup>8,9</sup>, Michel M. van den Heuvel<sup>10</sup>, Nicola Sverzellati<sup>5</sup>, Mathias Prokop<sup>1</sup>, Ugo Pastorino<sup>4</sup>, Cornelia M. Schaefer-Prokop<sup>1,11</sup> and Bram van Ginneken<sup>1,12</sup>

<sup>1</sup>Dept of Medical Imaging, Radboud University Medical Center, Nijmegen, The Netherlands. <sup>2</sup>Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, The Netherlands. <sup>3</sup>Dutch Expert Centre for Screening, Nijmegen, The Netherlands. <sup>4</sup>Unit of Thoracic Surgery, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy. <sup>5</sup>Section of Radiology, Unit of Surgical Sciences, Dept of Medicine and Surgery (DiMeC), University of Parma, Parma, Italy. <sup>6</sup>Dept of Biomedical Engineering and Physics, Amsterdam UMC – Location AMC, Amsterdam, The Netherlands. <sup>7</sup>Dept of Radiology and Nuclear Medicine, Amsterdam UMC – Location AMC, Amsterdam, The Netherlands. <sup>8</sup>Dept of Radiology, University Medical Center Utrecht, Utrecht, The Netherlands. <sup>9</sup>Utrecht University, Utrecht, The Netherlands. <sup>10</sup>Dept of Respiratory Diseases, Radboud University Medical Center, Nijmegen, The Netherlands. <sup>11</sup>Dept of Radiology, Meander Medisch Centrum, Amersfoort, The Netherlands. <sup>12</sup>Fraunhofer MEVIS, Bremen, Germany.

Corresponding author: Anton Schreuder ([antoniusschreuder@gmail.com](mailto:antoniusschreuder@gmail.com))



Shareable abstract (@ERSpublications)

**Lung cancer CT screening participants with a relatively low risk of lung cancer incidence and a high risk of competing death can be identified by applying two respective post-scan risk models, and in turn may benefit from other personalised trajectories** <https://bit.ly/2ZDe62K>

**Cite this article as:** Schreuder A, Jacobs C, Lessmann N, *et al.* Scan-based competing death risk model for re-evaluating lung cancer computed tomography screening eligibility. *Eur Respir J* 2022; 59: 2101613 [DOI: 10.1183/13993003.01613-2021].

Copyright ©The authors 2022.  
For reproduction rights and  
permissions contact  
[permissions@ersnet.org](mailto:permissions@ersnet.org)

Received: 7 June 2021  
Accepted: 17 Sept 2021

## Abstract

**Background** A baseline computed tomography (CT) scan for lung cancer (LC) screening may reveal information indicating that certain LC screening participants can be screened less, and instead require dedicated early cardiac and respiratory clinical input. We aimed to develop and validate competing death (CD) risk models using CT information to identify participants with a low LC risk and a high CD risk.

**Methods** Participant demographics and quantitative CT measures of LC, cardiovascular disease and chronic obstructive pulmonary disease were considered for deriving a logistic regression model for predicting 5-year CD risk using a sample from the National Lung Screening Trial (n=15 000). Multicentric Italian Lung Detection data were used to perform external validation (n=2287).

**Results** Our final CD model outperformed an external pre-scan model (CD Risk Assessment Tool) in both the derivation (area under the curve (AUC) 0.744 (95% CI 0.727–0.761) and 0.677 (95% CI 0.658–0.695), respectively) and validation cohorts (AUC 0.744 (95% CI 0.652–0.835) and 0.725 (95% CI 0.633–0.816), respectively). By also taking LC incidence risk into consideration, we suggested a risk threshold where a subgroup (6258/23 096 (27%)) was identified with a number needed to screen to detect one LC of 216 (*versus* 23 in the remainder of the cohort) and ratio of 5.41 CDs per LC case (*versus* 0.88). The respective values in the validation cohort subgroup (774/2287 (34%)) were 129 (*versus* 29) and 1.67 (*versus* 0.43).

**Conclusions** Evaluating both LC and CD risks post-scan may improve the efficiency of LC screening and facilitate the initiation of multidisciplinary trajectories among certain participants.

## Introduction

Various randomised controlled trials have demonstrated that lung cancer (LC) screening with low-dose computed tomography (CT) significantly reduces the number of LC deaths compared with chest radiography [1] or no screening [2, 3]. However, most deaths which occurred in these trials were among LC-free participants [1–3]. Even among participants who died of LC, most were reported to have other underlying causes of death. This indicates that preventing a LC death does not always lead to a gain in life-years [4, 5].

Screening eligibility is primarily based on two demographic predictors of LC incidence: age and smoking history [6, 7]. In addition, people with contraindications for curative LC treatments may be excluded. This is to ensure a sufficiently high detection rate of treatable LCs while avoiding potential harms caused by false-positive findings [8]. Various LC risk models exist specifically for determining screening participant eligibility among ever-smokers: the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial Model 2012 (PLCO<sub>M2012</sub>) and LC Risk Assessment Tool (LCRAT) are among the most renowned and best performing [9, 10].

After the baseline screening round, chest CT biomarkers can be used to improve prediction accuracy. Imaging features are especially beneficial for estimating nodule malignancy risk [11–17]. Moreover, quantitative CT (QCT) measures of cardiovascular disease (CVD) and chronic obstructive pulmonary disease (COPD) have also been validated as biomarkers in this setting [17–21]. However, determining eligibility for future screening rounds is based on the same criteria as before the first scan [6, 7]. Whereas post-scan LC incidence risk models have been developed to personalise screening intervals beyond 1 year [11, 12, 16, 17, 22], there are currently no guidelines for re-evaluating screening eligibility using CT information.

Regardless of their LC incidence risk, participants with a high risk of competing (*i.e.* non-LC) death (CD) may require multidisciplinary follow-up to benefit from screening. The ability to identify participants with a low risk of developing LC and a high risk of encountering a CD (low-LC-high-CD) may enable more personalised recommendations to continue benefiting those in need while reducing potential harms in others. Two primary objectives can be summarised for this study: 1) to develop a risk model combining demographic information and QCT measures for predicting CDs among LC screening participants, and 2) to demonstrate the potential implications of halting further participation of low-LC-high-CD screenees. A previously developed model which considers the same QCT measures from the baseline scan (*i.e.* of LC, CVD and COPD) was used to determine LC risk [17]. A secondary objective was to investigate whether a CD risk model combining both demographics and QCT measures as predictors is superior to only using one or the other.

## Methods

The “Scans and data”, “Models and variables” and “Dataset formation” subsections have been previously and more extensively described in SCHREUDER *et al.* [17]; more details are provided in the supplementary material.

### Scans and data

Scans and metadata from the National Lung Screening Trial (NLST) (ClinicalTrials.gov: NCT00047385) were used to derive models with permission from the National Cancer Institute (NCI) (Cancer Data Access System project: NLST-437) [1]. The NLST was the first and largest randomised controlled trial to demonstrate the effectiveness of CT screening in the reduction of LC death. Inclusion criteria were age 55–74 years, cigarette smoking history  $\geq 30$  pack-years and a smoking quit time within 15 years (if applicable). 26722 participants were assigned to the CT cohort and 26732 to the radiography cohort. Between 2002 and 2010, both cohorts underwent three annual screening rounds and were subsequently followed-up for 5 years (median follow-up time 6.5 years).

The Multicentric Italian Lung Detection (MILD) randomised controlled trial (ClinicalTrials.gov: NCT02837809) was used for external validation purposes [2]. MILD was the first trial to demonstrate the continued effectiveness of LC CT screening beyond the fifth year of screening. Inclusion criteria were age  $\geq 49$  years,  $\geq 20$  pack-years of smoking history and a smoking quit time no longer than 10 years (if applicable). 1190 participants were assigned to annual CT screening, 1186 to biennial CT screening and 1723 to no screening. Between 2005 and 2018, 93.5% of the participants were followed-up for at least 9 years; the median time until the last screening round was 6 years.

### Models and variables

Participant demographics and nodule CT features at baseline were directly provided by the NLST and MILD datasets. When multiple nodules were present in one scan, only the features from the nodule with the longest diameter were considered for the models. In the NLST, causes of death were determined by the end-point verification process (or death certificate if unavailable). Causes of death for MILD were obtained exclusively from death certificates. Only underlying cause of death was considered for this study, either LC death or CD.

QCT measures of CVD and COPD were extracted from the baseline CT scans using validated computer algorithms. These included calcium volume and mean calcium density in the coronary arteries (combined), mitral valve, aortic valve and transthoracic aorta [19], emphysema score (percentage of lung voxels  $< -950$  HU after resampling the CT images to 3 mm slice thickness, normalisation and bullae analysis) [23], and a reference parameter for measuring bronchial wall thickness called Pi10 (the square root of the airway wall area for a theoretical 10 mm lumen perimeter airway derived using the linear regression of the square root of segmented wall areas against the lumen perimeter extracted from the complete segmented airway tree) [20].

Three models for predicting the 5-year risk of CD were derived: a model containing only self-reported demographic information ( $CD_{\text{survey}}$ ), a model containing only age, sex and CT information ( $CD_{\text{CT}}$ ), and a final model considering all variables from the previous two models ( $CD_{\text{final}}$ ). The previously described  $LCi_{\text{final}}$  model was used to calculate the 5-year LC incidence probability (supplementary equations S1) [17]; the model had been calibrated to each cohort. Table 1 lists all variables considered.

#### Dataset formation

Subject exclusion criteria were the lack of a baseline CT scan, a baseline scan with a slice thickness  $>2.5$  mm, and missing data on LC incidence, death status and times of event. The derivation cohort consisted of 15000 unique subjects from the NLST CT cohort, which was the maximum allowed data extraction permitted by the NCI (supplementary figure S1). This included all participants who were diagnosed with LC or died within the trial period ( $n=2106$ ) and a random sample of those who did not encounter either of the events ( $n=12894$ ); this latter group was resampled without replacement ( $n=8096$ ), ultimately forming a derivation cohort of 23096 subjects. This method for weighted analysis was performed to simulate the original NLST cohort, thereby preventing an overestimation of risk.

All 2287 subjects from the MILD CT cohort formed the validation cohort. Multiple imputation was used to create plausible values for other missing data (supplementary methods). Missing QCT measures were considered an exclusion criterion for the derivation cohort, but not the validation cohort, and were given the respective median values from the MILD dataset ( $n=24$  for QCT measures of CVD and  $n=132$  for Pi10). In addition, a MILD subgroup of NLST-eligible participants was formed, namely those 55–74 years of age,  $\geq 30$  pack-years of smoking intensity and no more than 15 years since smoking cessation (if applicable) [1].

#### Statistical analysis

All statistical analyses were performed in R version 3.4.3 ([www.r-project.org](http://www.r-project.org)). The three parsimonious risk prediction models for CD were derived using logistic regression. Variables were included in the model if the level of significance was  $<0.20$  (backward elimination) [24]. R function “mfp” (mfp package) was used to select factorial polynomials for continuous variables (level of significance  $<0.05$ ) [25]. Second-degree polynomials were considered for the longest nodule diameter; only first-degree polynomials were considered for the remaining variables.

Area under the receiver operating characteristic curve (AUC) values with 95% confidence intervals were calculated for our derived models ( $CD_{\text{survey}}$ ,  $CD_{\text{CT}}$  and  $CD_{\text{final}}$ ) and the previously described CD Risk Assessment Tool (CDRAT) [10]. Internal validation was performed for each model with 1000 bootstraps to assess overfitting in the form of optimism (the difference between the average bootstrap sample AUC and original sample AUC). The values were compared with 500 bootstraps to test whether the discriminative performances were significantly different ( $p<0.05$ ). Calibration was visualised in the form of calibration plots. The models were externally validated on the full MILD cohort.

To demonstrate practical application, the cohorts were stratified into quintiles (five equally sized groups) based on 5-year risk of LC incidence ( $LCi_{\text{final}}$ ) and competing death ( $CD_{\text{final}}$ ). The quintiles were used to form  $5\times 5$  contingency tables showing the distribution of LC diagnoses and CDs within 5 years of follow-up. Based on visual findings, we suggest criteria to select low-LC-high-CD participants. The potential consequences of applying these selection criteria are summarised by calculating the number needed to screen (NNS) to detect one LC and the ratio of CDs per LC diagnosis. The same risk probability thresholds were applied to the full and NLST-eligible MILD cohorts for external validation. This analysis was repeated using external pre-scan risk models for LC risk (*i.e.*  $PLCO_{M2012}$  and LCRAT) and CD risk (*i.e.* CDRAT) [10, 26], using risk probability thresholds based on model-dependent quintiles. McNemar’s test was used to determine whether the frequencies between the different low-LC-high-CD groups were statistically different [27].

TABLE 1 Distribution of variables by 5-year competing death (CD) event or not

|   | Derivation cohort (NLST) |                  | Validation cohort (MILD) |                      |
|---|--------------------------|------------------|--------------------------|----------------------|
|   | CD (n=800)               | No CD (n=22 296) | CD (n=33)                | No CD (n=2254)       |
| <b>Patient characteristics</b>                      |                          |                  |                          |                      |
| Age, years  | 63.6±5.5                 | 61.3±5.0         | 62.8±5.8                 | 57.5±5.9             |
| Female  | 223 (27.9)               | 9286 (41.6)      | 8 (24.2)                 | 717 (31.8)           |
| Race or ethnicity                                   |                          |                  |                          |                      |
| White   | 714 (89.3)               | 19 963 (89.5)    | 33 (100)                 | 2250 (99.8)          |
| Black   | 51 (6.4)                 | 1000 (4.5)       | 0                        | 0                    |
| Asian   | 6 (0.8)                  | 542 (2.4)        | 0                        | 2 (0.1)              |
| Hispanic  | 7 (0.9)                  | 335 (1.5)        | 0                        | 1 (0.0)              |
| Mixed or other                                      | 22 (2.8)                 | 456 (2.0)        | 0                        | 1 (0.0)              |
| Educational level, 0–5 <sup>#</sup>                 | 3.6±1.7                  | 3.7±1.6          | 1.7±1.5 <sup>f</sup>     | 1.4±1.4 <sup>f</sup> |
| BMI, kg·m <sup>-2</sup>                             | 27.1 (24.0–30.9)         | 27.2 (24.4–30.5) | 26.1 (22.4–30.5)         | 25.7 (23.5–28.4)     |
| Current smoker                                      | 455 (56.9)               | 10 564 (47.4)    | 22 (66.7)                | 1544 (68.5)          |
| Smoking intensity, pack-years                       | 54 (43–76)               | 48 (39–66)       | 45 (34–59)               | 39 (32–51)           |
| Smoking duration, years                             | 43.2±7.4                 | 39.7±7.3         | 43.1±7.1                 | 38.1±6.7             |
| Smoking quit time, years <sup>†</sup>               | 7 (3–12)                 | 8 (4–12)         | 4 (2–6)                  | 5 (3–8)              |
| LC in family <sup>‡</sup>                           |                          |                  |                          |                      |
| 1   | 141 (17.6)               | 4088 (18.3)      | 3 (9.1)                  | 550 (24.4)           |
| ≥2  | 30 (3.8)                 | 712 (3.2)        | NA                       | NA                   |
| Work asbestos                                       | 58 (7.3)                 | 1017 (4.6)       | 2 (6.1)                  | 195 (8.7)            |
| COPD diagnosis <sup>§</sup>                         | 217 (27.1)               | 3943 (17.7)      | 6 (18.2)                 | 266 (11.8)           |
| Asthma diagnosis                                    | 102 (12.8)               | 2158 (9.7)       | 3 (9.1)                  | 143 (6.3)            |
| Diabetes diagnosis                                  | 145 (18.1)               | 2062 (9.2)       | 4 (12.1)                 | 132 (5.9)            |
| Heart disease diagnosis                             | 195 (24.4)               | 2707 (12.1)      | 4 (12.1)                 | 277 (12.3)           |
| Hypertension diagnosis                              | 363 (45.4)               | 7665 (34.4)      | 14 (42.4)                | 619 (27.5)           |
| Stroke diagnosis                                    | 58 (7.3)                 | 565 (2.5)        | 2 (6.1)                  | 20 (0.9)             |
| <b>Nodule CT features</b>                           |                          |                  |                          |                      |
| Nodule attenuation                                  |                          |                  |                          |                      |
| No nodule   | 600 (75.0)               | 16 296 (73.1)    | 18 (54.5)                | 980 (43.5)           |
| Solid   | 152 (19.0)               | 4636 (20.8)      | 11 (33.3)                | 1004 (44.5)          |
| Part solid  | 14 (1.8)                 | 359 (1.6)        | 0                        | 61 (2.7)             |
| Nonsolid  | 34 (4.3)                 | 1005 (4.5)       | 4 (12.1)                 | 209 (9.3)            |
| Longest diameter, mm <sup>¶</sup>                   | 7 (5–11)                 | 6 (5–9)          | 4.6 (3.7–8.5)            | 4.9 (3.5–7.4)        |
| Perpendicular diameter, mm <sup>¶</sup>             | 5 (4–8)                  | 5 (4–7)          | 3.5 (3.0–5.1)            | 3.9 (2.8–5.8)        |
| Nodule in upper lobe <sup>¶</sup>                   | 88 (41.7)                | 2496 (39.1)      | 7 (21.2)                 | 370 (16.4)           |
| Nodule spiculation <sup>¶</sup>                     | 30 (14.2)                | 789 (12.4)       | NA                       | NA                   |
| Nodule count <sup>¶</sup>                           | 1 (1–2)                  | 1 (1–2)          | 1 (1–2)                  | 1 (1–2)              |
| <b>Quantitative CT measures of CVD</b>              |                          |                  |                          |                      |
| Coronary calcium volume, mm <sup>3</sup>            | 188 (36–788)             | 48 (2–267)       | 114 (21–466)             | 23 (0–154)           |
| Coronary mean calcium density, HU                   | 226 (190–264)            | 207 (141–251)    | 286 (227–314)            | 255 (0–311)          |
| Transthoracic aorta calcium volume, mm <sup>3</sup> | 1134 (314–2894)          | 403 (89–1282)    | 948 (200–2992)           | 200 (45–694)         |
| Transthoracic aorta mean calcium density, HU        | 326 (277–377)            | 311 (251–378)    | 418 (375–451)            | 434 (363–523)        |
| Mitral valve calcium volume, mm <sup>3</sup>        | 0 (0–18)                 | 0 (0–0)          | 0 (0–2)                  | 0 (0–0)              |
| Mitral valve mean calcium density, HU               | 0 (0–206)                | 0 (0–0)          | 0 (0–203)                | 0 (0–0)              |
| Aortic valve calcium volume, mm <sup>3</sup>        | 0 (0–20)                 | 0 (0–0)          | 0 (0–33)                 | 0 (0–0)              |
| Aortic valve mean calcium density, HU               | 0 (0–181)                | 0 (0–0)          | 0 (0–241)                | 0 (0–0)              |
| <b>Quantitative CT measures of COPD</b>             |                          |                  |                          |                      |
| Total lung volume, L                                | 5.6 (4.7–6.7)            | 5.4 (4.5–6.4)    | 6.2 (5.2–6.9)            | 5.9 (5.1–6.8)        |
| Mean lung density, HU                               | –834 (–857––808)         | –839 (–858––815) | –845 (–872––834)         | –846 (–861––828)     |
| Emphysema score                                     | 0.39 (0.08–1.77)         | 0.24 (0.05–1.12) | 0.08 (0.00–1.22)         | 0.03 (0.00–0.17)     |
| Pi10  | 3.0 (2.5–3.6)            | 2.8 (2.3–3.3)    | 2.5 (2.2–3.0)            | 2.4 (2.2–2.6)        |

Data are presented as mean±SD, n (%) or median (interquartile range). NLST: National Lung Screening Trial; MILD: Multicentric Italian Lung Detection; BMI: body mass index; LC: lung cancer; COPD: chronic obstructive pulmonary disease; CT: computed tomography; NA: not applicable; CVD: cardiovascular disease; Pi10: measure of bronchial wall thickness. <sup>#</sup>: a categorical variable applied as a continuous variable, where 0=did not complete high school, 1=high school graduate, 2=post-high school training but no college, 3=some college, 4=bachelor's degree and 5=graduate school or higher; <sup>¶</sup>: of those applicable (regarding nodule features, applies to only the nodule with the longest diameter); <sup>†</sup>: number of first-degree family members diagnosed with LC (a value of "2" was given when two or more family members were diagnosed); <sup>§</sup>: includes prior diagnosis of COPD, emphysema and/or chronic bronchitis; <sup>f</sup>: on a scale of 0–4. Inter-cohort statistics (p-values and effect sizes) are reported in Schreuder *et al.* [17].

## Results

### Study participants

Within 5 years of follow-up, 756 LC diagnoses (3.3%) and 800 CDs (3.5%) were reported in the derivation cohort (n=23096); 22 participants encountered both outcomes. The respective numbers in the validation cohort (n=2287) were 59 (2.6%) and 33 (1.4%), with two participants overlapping. On average, the NNS in the derivation and validation cohorts was 31 (23096/756) and 39 (2287/59), respectively. The respective ratios of CDs per LC diagnosed within 5 years of follow-up were 1.06 (800/756) and 0.56 (33/59). Descriptive statistics are summarised in table 1.

### CD risk prediction

Three CD risk models were derived; the variables and coefficients of CD<sub>final</sub>, CD<sub>survey</sub> and CD<sub>CT</sub> are reported in table 2, and supplementary tables S1 and S2, respectively. The AUC of CD<sub>final</sub> in the derivation cohort was 0.744 (95% CI 0.727–0.761), significantly greater than that of CD<sub>survey</sub> (0.707, 95% CI 0.689–0.725) and CD<sub>CT</sub> (0.719, 95% CI 0.701–0.737) (p<0.001) (supplementary figure S2). Internal validation revealed an optimism no greater than 0.006 (supplementary table S3). The AUC of the CDRAT model was 0.677 (95% CI 0.658–0.695), significantly inferior to that of CD<sub>survey</sub> (p<0.001).

External validation resulted in AUCs of 0.744 (95% CI 0.652–0.835), 0.721 (95% CI 0.627–0.815), 0.756 (95% CI 0.667–0.844) and 0.725 (95% CI 0.633–0.816) for CD<sub>final</sub>, CD<sub>survey</sub>, CD<sub>CT</sub> and CDRAT, respectively (no statistically significant differences) (supplementary figure S3). The calibration curves of CD<sub>final</sub> in the derivation and validation cohort are shown in supplementary figures S4 and S5, respectively. Note that the deviation of the calibration curve from the diagonal in the validation cohort is caused by one outlier case (5-year CD risk probability 19.7%).

Decision curve analyses of the CD models are available in supplementary figures S6 and S7 [28].

**TABLE 2** Final competing death risk model (CD<sub>final</sub>)

| Variable   | β coefficient | OR (95% CI)         | p-value |
|--|---------------|---------------------|---------|
| Model intercept                                    | −3.73776      | 0.02 (0.01–0.10)    | <0.001  |
| Age, per year                                      | 0.02399       | 1.02 (1.01–1.04)    | 0.006   |
| Sex, female  | −0.54489      | 0.58 (0.49–0.69)    | <0.001  |
| White race, reference                              | NA            | NA                  | NA      |
| Black race, yes                                    | 0.39969       | 1.49 (1.08–2.01)    | 0.011   |
| Asian race, yes                                    | −1.24615      | 0.29 (0.11–0.60)    | 0.003   |
| Hispanic race, yes                                 | −0.36998      | 0.69 (0.29–1.37)    | 0.341   |
| Mixed or other race, yes                           | 0.34995       | 1.42 (0.88–2.16)    | 0.123   |
| BMI, per kg·m <sup>−2</sup> #                      | 3.35128       | 28.54 (5.52–143.16) | <0.001  |
| Current smoker, yes                                | 0.27488       | 1.32 (1.11–1.56)    | 0.002   |
| Smoking duration, per year#                        | −0.04392      | 0.96 (0.92–0.99)    | 0.016   |
| Hypertension diagnosis, yes                        | 0.12417       | 1.13 (0.97–1.32)    | 0.118   |
| Diabetes diagnosis, yes                            | 0.45764       | 1.58 (1.29–1.93)    | <0.001  |
| Heart disease diagnosis, yes                       | 0.20551       | 1.23 (1.01–1.48)    | 0.034   |
| Stroke diagnosis, yes                              | 0.55247       | 1.74 (1.28–2.32)    | <0.001  |
| Asthma diagnosis, yes                              | 0.33797       | 1.40 (1.11–1.75)    | 0.004   |
| COPD diagnosis, yes                                | 0.25723       | 1.29 (1.08–1.54)    | 0.004   |
| Emphysema score, per point                         | 0.05331       | 1.05 (1.04–1.07)    | <0.001  |
| Bronchial wall thickness (Pi10), per point         | 0.16154       | 1.18 (1.06–1.30)    | 0.001   |
| Mean lung density, per HU#                         | −1.83652      | 0.16 (0.08–0.33)    | <0.001  |
| Aorta calcium volume, per mm <sup>3</sup> #        | 0.18081       | 1.20 (1.14–1.26)    | <0.001  |
| Aorta calcium mean density, per HU#                | −1.53207      | 0.22 (0.09–0.50)    | 0.001   |
| Coronary calcium volume, per mm <sup>3</sup> #     | 0.14995       | 1.16 (1.05–1.28)    | 0.003   |
| Mitral valve calcium volume, per mm <sup>3</sup> # | 0.06940       | 1.07 (1.05–1.10)    | <0.001  |

BMI: body mass index; COPD: chronic obstructive pulmonary disease; NA: not applicable. To calculate the 5-year risk probability of lung cancer incidence, first find the sum of the products of each (transformed) variable and their respective β coefficient to obtain the linear predictor, then insert the value into the following equation:  $1/(1+e^{-(\text{linear predictor})})$ . #: transformed accordingly as: (BMI/10)<sup>−2</sup>, (smoking duration/100)<sup>−2</sup>, ((mean lung density+1000)/100)<sup>−1</sup>, ln((aorta calcium volume+0.1)/1000), (aorta calcium mean density/1000), (coronary calcium volume/1000) and ln((mitral valve calcium volume+0.1)/100).

### **Stratification by LC risk and CD risk**

Figure 1 is a collection of three-dimensional column charts and 2×2 contingency tables divided into cells based on risk quintiles according to  $CD_{final}$  (vertical axis) and  $LCi_{final}$  (horizontal axis). The cut-off values are reported in supplementary table S4. Contingency tables for the validation cohort are given in supplementary table S5 and supplementary figure S8.

Based on the 5-year risk estimate, we determined visually that participants with a  $LCi_{final}$  risk  $\leq 0.79\%$ ,  $LCi_{final}$  risk  $\leq 1.38\%$  and  $CD_{final}$  risk  $> 2.93\%$ , or  $LCi_{final}$  risk  $\leq 2.18\%$  and  $CD_{final}$  risk  $> 4.92\%$ , had a relatively high ratio (generally  $> 3$ ) of CDs per LC diagnosis. 27% (6258/23096) of the derivation cohort fit these criteria and were stratified into the low-LC-high-CD group (table 3). Four percent (29/756) of the LC cases and 20% (157/800) of the CDs occurred in this group; the NNS was 216 (cohort average 31) and 5.41 CDs occurred per LC diagnosis (cohort average 1.06).

The same risk probability thresholds were applied for external validation (table 3). 34% (774/2287) of MILD participants fell into the low-LC-high-CD group, consisting of 10% (6/59) of the LC cases and 30% (10/33) of the CDs. The NNS was 129 (full validation cohort average 39) and 1.67 CDs occurred per LC diagnosis (average 0.56). When only considering NLST-eligible MILD participants, the resulting proportions, ratios and NNS were closer to those from the derivation cohort.

The performance of three external pre-scan models was also assessed using the same methods, first by combining LCRAT with CDRAT and then by combining  $PLCO_{M2012}$  with CDRAT (table 3). Note that the stratification criteria were dependent on the models used and their risk probability distributions in the derivation cohort. Compared with the post-scan models, significantly more LC cases ( $p < 0.001$  for both LCRAT and  $PLCO_{M2012}$ ) and fewer CDs ( $p = 0.048$  for LCRAT and  $p = 0.012$  for  $PLCO_{M2012}$ ) were stratified to the low-LC-high-CD groups based on pre-scan model risk. A similar trend was found in the validation cohort, but the differences between the pre- and post-test models were not statistically significant. In the NLST-eligible validation cohort, the external models resulted in a lower ratio of CDs per LC diagnosis in the low-LC-high-CD group than in the high-LC-low-CD group.

The outcomes of using  $CD_{CT}$ ,  $CD_{survey}$  or CDRAT for stratifying a low-LC-high-CD group given the same LC risk thresholds ( $LCi_{final}$ ) are available for comparison in supplementary table S6.

## **Discussion**

### **CD risk prediction**

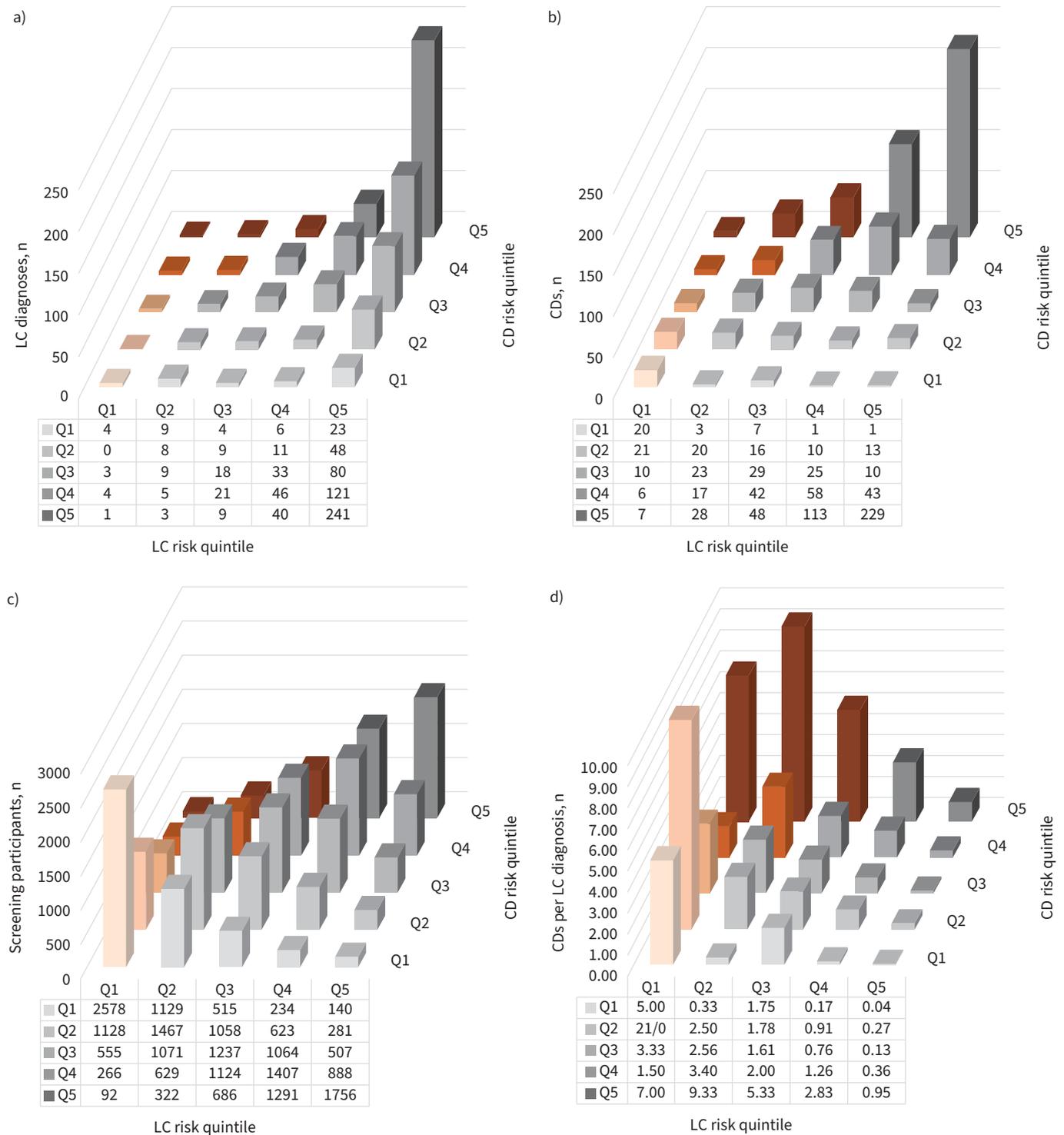
Three 5-year CD risk prediction models were derived, including only self-reported patient demographics ( $CD_{survey}$ ), only QCT measures ( $CD_{CT}$ ) or both ( $CD_{final}$ ).  $CD_{CT}$  and  $CD_{final}$  are the first models using the baseline CT scan for this purpose.  $CD_{survey}$  and the external model CDRAT were used to quantify the added value of CT information in addition to pre-scan information (supplementary figures S2 and S3) [10].

The discriminative performance of  $CD_{final}$  was significantly superior to the other models in the derivation cohort (supplementary figure S2). Due to the small differences in AUC and smaller sample size of MILD, no significant differences were determined between the models in the validation cohort (supplementary figure S3). However, the trend remained that CT-based models ( $CD_{CT}$  and  $CD_{final}$ : AUC 0.756 and 0.744, respectively) had a higher accuracy than pre-scan models (CDRAT and  $CD_{survey}$ : AUC 0.725 and 0.721, respectively). This suggests that a model based exclusively on QCT measures ( $CD_{CT}$ ) may be a viable option for automatically calculating risk scores in a high-risk population.

### **Assessment of model predictors**

The most notable observation of the variables included in  $CD_{final}$  is the lack of nodule CT features. Despite there being a seemingly high correlation between the risks of LC and CD (figure 1 and supplementary figure S8), nodule information is not sufficiently distinctive compared with the contribution of other risk factors of non-LC causes of death. In  $CD_{CT}$ , only the presence of a solid opacity as the largest nodule was included; this had a negative  $\beta$  coefficient, suggesting that participants with a nodule were more likely to encounter a LC death instead of a CD. In turn, QCT measures of CVD and COPD significantly improved the accuracy of  $CD_{final}$  and  $CD_{CT}$ . It had been previously shown that these QCT measures consistently contributed to the prediction of CVD death, COPD death, LC death and LC incidence [17].

Regarding measures of cardiovascular calcifications, we observed that a lower aorta mean calcium density was associated with a greater CD risk. A lower plaque density may indicate instability in an elderly population and therefore be more likely to be the cause of death [17, 29, 30]. This population with overall high CVD risk may represent a niche for specific risk stratification as opposed to the Agatston score used



**FIGURE 1** Outcomes per lung cancer (LC) and competing death (CD) risk quintile (derivation cohort). Three-dimensional column charts of **a)** LC diagnoses, **b)** CDs, **c)** screening participants and **d)** CDs per LC diagnosis (*y*-axis truncated at 10). Each CD risk quintile is differently shaded, where a darker shade corresponds to a higher risk. The columns are divided into grey and orange columns indicating the suggested separation of screening participants into a group which should continue to be screened (grey) and a group with a relatively low LC risk and high risk of CD (orange). Quintile 1 (Q1) represents the lowest risk quintile and quintile 5 (Q5) represents the highest risk quintile.

in populations with much broader risk ranges [31]. Conversely, a positive association was observed between mitral valve mean calcium density and CD risk (CD<sub>CT</sub>), which suggests that the pathophysiological mechanism of plaques is location dependent.

TABLE 3 Clinical outcomes of risk stratification

| Risk models used                              | Risk group <sup>#</sup>         | Participants | 5-year LC diagnoses | 5-year CDs | NNS to detect 1 LC | CDs per LC diagnosis |
|---|---------------------------------|--------------|---------------------|------------|--------------------|----------------------|
| <b>Derivation cohort (NLST)</b>               |                                 |              |                     |            |                    |                      |
| LC <sub>final</sub> and CD <sub>final</sub>   | High-LC-low-CD                  | 16 838 (73)  | 727 (96)            | 643 (80)   | 23                 | 0.88                 |
|   | Low-LC-high-CD                  | 6258 (27)    | 29 (4)              | 157 (20)   | 216                | 5.41                 |
| LCRAT and CDRAT                               | High-LC-low-CD                  | 17 335 (75)  | 677 (90)            | 668 (84)   | 26                 | 0.99                 |
|   | Low-LC-high-CD                  | 5761 (25)    | 79 (10)             | 132 (16)   | 73                 | 1.67                 |
| PLCO <sub>M2012</sub> and CDRAT               | High-LC-low-CD                  | 17 277 (75)  | 686 (91)            | 675 (84)   | 25                 | 0.99                 |
|   | Low-LC-high-CD                  | 5819 (25)    | 70 (11)             | 125 (16)   | 83                 | 1.79                 |
|   | Full derivation cohort          | 23 096 (100) | 756 (100)           | 800 (100)  | 31                 | 1.06                 |
| <b>Validation cohort (MILD)</b>               |                                 |              |                     |            |                    |                      |
| LC <sub>final</sub> and CD <sub>final</sub>   | High-LC-low-CD                  | 1513 (66)    | 53 (90)             | 23 (70)    | 29                 | 0.43                 |
|   | Low-LC-high-CD                  | 774 (34)     | 6 (10)              | 10 (30)    | 129                | 1.67                 |
| LCRAT and CDRAT                               | High-LC-low-CD                  | 1511 (66)    | 48 (81)             | 22 (67)    | 31                 | 0.46                 |
|   | Low-LC-high-CD                  | 776 (34)     | 11 (19)             | 11 (33)    | 71                 | 1.00                 |
| PLCO <sub>M2012</sub> and CDRAT               | High-LC-low-CD                  | 1568 (69)    | 51 (86)             | 28 (85)    | 31                 | 0.55                 |
|   | Low-LC-high-CD                  | 719 (31)     | 8 (14)              | 5 (15)     | 90                 | 0.63                 |
|   | Full validation cohort          | 2287 (100)   | 59 (100)            | 33 (100)   | 39                 | 0.56                 |
| <b>NLST-eligible validation cohort (MILD)</b> |                                 |              |                     |            |                    |                      |
| LC <sub>final</sub> and CD <sub>final</sub>   | High-LC-low-CD                  | 980 (80)     | 44 (94)             | 18 (72)    | 22                 | 0.41                 |
|   | Low-LC-high-CD                  | 245 (20)     | 3 (6)               | 7 (28)     | 82                 | 2.33                 |
| LCRAT and CDRAT                               | High-LC-low-CD                  | 995 (81)     | 39 (83)             | 21 (84)    | 26                 | 0.54                 |
|   | Low-LC-high-CD                  | 230 (19)     | 8 (17)              | 4 (16)     | 29                 | 0.50                 |
| PLCO <sub>M2012</sub> and CDRAT               | High-LC-low-CD                  | 1088 (89)    | 43 (91)             | 25 (100)   | 25                 | 0.58                 |
|   | Low-LC-high-CD                  | 137 (11)     | 4 (9)               | 0 (0)      | 34                 | 0.00                 |
|   | NLST-eligible validation cohort | 1225 (100)   | 47 (100)            | 25 (100)   | 26                 | 0.53                 |

Data are presented as n (%) or n. LC: lung cancer; CD: competing death; NNS: number needed to screen; NLST: National Lung Screening Trial; LC<sub>final</sub>: final LC incidence model; CD<sub>final</sub>: final CD model; LCRAT: LC Risk Assessment Tool; CDRAT: CD Risk Assessment Tool; PLCO<sub>M2012</sub>: Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial Model 2012; MILD: Multicentric Italian Lung Detection. #: high-LC-low-CD: group with LC risk >60th percentile, LC risk >40th percentile and CD risk ≤80th percentile, or LC risk >20th percentile and CD risk ≤60th percentile; low-LC-high-CD: group with LC risk ≤20th percentile, LC risk ≤40th percentile and CD risk >60th percentile, or LC risk ≤60th percentile and CD risk >80th percentile.

No other unexpected relationships between the variables and CD were observed. Note that no tests for multicollinearity or interactions between variables were performed for this study. Therefore, causative relations between variables and outcomes should not be deduced based solely on these findings.

#### Stratification by LC risk and CD risk

The NELSON LC screening trial results suggested that fixed screening intervals of >2 years are contraindicated [32, 33]. At the same time, the idea that screening intervals can be lengthened among participants determined to have a post-scan low LC risk is gaining momentum [2, 11, 12, 17, 22]. SCHREUDER *et al.* [17] proposed screening intervals extending past 2 years among participants with a sufficiently low LC risk. However, an approach which focuses solely on LC risk may not be optimal because the potential benefits of early LC detection may be humbled by underlying competing morbidities. Other pre-scan risk models have applied approaches based on the likelihood to survive LC screening, namely by predicting LC death risk [10, 34] or life-years gained [4, 5].

Based on this idea that low-LC-high-CD participants are less likely to benefit from LC screening, we propose to consider multidisciplinary follow-up among a small group of individuals. This may involve proactive trajectories according to cardiac and respiratory disease management guidelines for participants who are not yet under the care of a physician. An example of opportunistic risk assessment is given by the Manchester Lung Screening Pilot, which found that 47% of the participants at high risk of CVD (QRISK2 score ≥10%) were not taking lipid-lowering medication, which is indicated for primary prevention [35].

Additionally, a longer screening interval of up to 5 years can be considered based on post-scan risk prediction. In the derivation cohort, around a fourth (27% (6258/23096)) of the screenees would have received a counter-indication to continue participating. This would be at the expense of delaying the LC

diagnosis in 4% (29/756) of the participants who would develop LC. We considered this a pessimistic estimate because the analysis did not consider participants undergoing additional follow-up CTs in response to suspicious findings in the baseline scan (4/29 participants were diagnosed within 1 year of the baseline scan); this group also includes participants which were both diagnosed with LC and encountered a CD (1/29). In exchange, 6229 people may have hypothetically avoided four annual screening rounds which would not have resulted in a LC diagnosis. 2.5% (157/6258) of low-LC-high-CD participants encountered CDs within 5 years, equivalent to 20% of all CDs (157/800). We caution that advocating no further screening may negate some of the positive psychological benefits, *e.g.* interest in lifestyle advice [36].

Using the same risk probability thresholds, a similar trend occurred in the validation cohort. A greater proportion of participants fell within the low-LC-high-CD group (34% (774/2287)) because the median 5-year calibrated LC incidence risk was lower than in the derivation cohort (1.28% and 1.73%, respectively); this was despite the median CD risk also being lower (1.40% and 2.36%, respectively). This would have resulted in a greater proportion of both LC diagnoses (10% (6/59)) and CDs (30% (10/33)) within the low-LC-high-CD group.

Compared with pre-scan models, using  $LC_{i_{final}}$  and  $CD_{final}$  resulted in higher values for both the NNS and CDs per LC diagnosis in the low-LC-high-CD group. The differences were not statistically significant in the validation cohort despite the similar trend. The risk thresholds used for stratification in our study were merely suggestive; we encourage adjusting the thresholds to satisfy each screening programme's aims and values. At certain thresholds, a pre-scan CD model may even perform equivalently to a post-scan CD model.

### Limitations

The primary limitation is that the external validation cohort had a low number of events of interest (less than 100). This was reflected in the wide confidence intervals. It was therefore also not considered useful to calibrate or retrain the model to MILD data. Once more LC screening CT data become available, more extensive model validation is recommended before considering implementation in practice.

Not having used Cox proportional hazards regression to model CD also introduces possible biases related to censoring (*i.e.* loss to follow-up) and competing risk from LC-specific mortality. However, logistic regression does not need to assume proportional hazards and is easier to interpret and implement as a risk model, which was sufficient for the purpose of this study [37, 38]. Although the coefficients of logistic regression may be a bit inflated compared with those from Cox regression, it is the predictive performance rather than the association between the predictors and outcome which is relevant. The analyses also ignored the fact that most participants who died of LC were reported to have multiple secondary causes of death. The risk score was arbitrarily restricted to a 5-year post-scan period to limit the loss to follow-up while being considerably longer than the standard 1-year screening interval.

Another limitation is that nodule diameter was measured manually in the NLST (no volumetric data were available) and only the data from the nodule with the longest diameter were considered when multiple nodules were detected. Volume (or mean diameter) obtained by semiautomatic software would offer better discrimination [39, 40]. Whereas nodule size is the most predictive variable for malignancy (given a single scan), this risk can be further modified based on other features, mainly nodule type, location and presence of spiculation [11, 13, 17]. Note that nodule spiculation was not recorded by MILD. Also, QCT measures automatically obtained from scans with slice thickness >1 mm (namely from the NLST) may not be reliable (*i.e.* emphysema score, Pi10 and calcium scores). Another possible issue with the measurement of CT calcium volume and density is that the scans were not ECG-gated, but prior studies indicate a strong predictive value [41] and high concordance with gated calcium scores [42].

### Future directions

Additional diagnostic interventions (usually in the form of follow-up CTs) are recommended for nodules with an indeterminate malignancy risk [43, 44]. In current screening practice, this additional work-up does not affect the timing of the subsequent annual screening rounds. We hypothesise that it would be of added value to know the outcome of additional diagnostic tests before deciding on post-scan screening eligibility (or screening interval length). Besides nodule growth, the availability of follow-up scans would enable the estimation of CVD and COPD progression rates. An alternative would be to manage nodules independently of other decisions, *e.g.* a low-LC-high-CD participant with an indeterminate nodule would still be followed-up according to nodule guideline recommendations while being ineligible for further screening. Regardless, future studies should simulate decision trees at multiple time-points while taking time-varying risk factors into consideration.

### Conclusions

We derived 5-year CD risk models using either self-reported patient characteristics ( $CD_{\text{survey}}$ ), chest CT image biomarkers ( $CD_{\text{CT}}$ ) or both ( $CD_{\text{final}}$ ). QCT measures of CVD and COPD were included in the CT-based models; pulmonary nodule morphological features were not found to be significant predictors. CT information provides an added value to the AUC of at least 2 percentage points. In a high-risk screening population, there may be little or no added value of patient demographics for predicting CD when QCT measures have been extracted: a CT scan alone may elucidate personalised susceptibility to smoke damage, ageing and other risk factors for CD.

By calculating both post-scan LC incidence risk ( $LC_{\text{final}}$ ) and CD risk ( $CD_{\text{final}}$ ), a group of participants with a relatively low risk of the former and high risk of the latter can be identified (low-LC-high-CD). This means that the baseline scan can be used to help identify participants who may benefit from multidisciplinary action and can safely be recommended longer screening intervals. Using our suggested criteria for contraindicating LC screening participation within the next 5 years may reduce the number of screenees by approximately a fourth, of whom more than 200 would need to continue participating to detect one LC. Valuable healthcare resources could simultaneously be redirected towards the prevention of CDs among low-LC-high-CD participants.

**Acknowledgements:** The authors thank Gabriel Humpire Mamani (Radboud University Medical Center, Nijmegen, The Netherlands), Jean-Paul Charbonnier and Leticia Gallardo-Estrella (Thirona, Nijmegen, The Netherlands) for their technical support in obtaining quantitative CT measures, and Claudio Jacomelli and Frederica Sabia (Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy) for extracting the requested Multicentric Italian Lung Detection (MILD) trial data. The authors also thank the MILD research teams for access to MILD data and the National Cancer Institute (NCI) for access to the NCI's data collected by the National Lung Screening Trial (NLST) under project number NLST-437. The statements contained herein are solely those of the authors and do not represent or imply concurrence or endorsement by the NCI.

An online risk calculator for the models described can be accessed at: <https://docs.google.com/spreadsheets/d/1IU-UH1mxOml-O--sNo8lhAgu2WLzRkcOQBCEm4rLdb4/edit?usp=sharing>

**Author contributions:** A. Schreuder: conceived and designed the analysis, performed the analysis, and wrote the paper. C. Jacobs: conceived the analysis, supervised the analysis, collected the data, contributed data and analysis tools, and critically appraised the paper. N. Lessmann: conceived the analysis, collected the data, contributed data and analysis tools, and critically appraised the paper. M.J.M. Broeders: conceived the analysis, supervised the analysis and critically appraised the paper. M. Silva: conceived the analysis, collected the data, contributed data, and critically appraised the paper. I. Išgum: conceived the analysis, contributed analysis tools and critically appraised the paper. P.A. de Jong: conceived the analysis, contributed analysis tools and critically appraised the paper. M.M. van den Heuvel: conceived the analysis and critically appraised the paper. N. Sverzellati: conceived the analysis, collected the data, contributed data and critically appraised the paper. M. Prokop: conceived the analysis and critically appraised the paper. U. Pastorino: conceived the analysis, collected the data, contributed data and critically appraised the paper. C.M. Schaefer-Prokop: conceived the analysis and critically appraised the paper. B. van Ginneken: conceived the analysis, supervised the analysis, contributed analysis tools and critically appraised the paper.

**Conflict of interest:** A. Schreuder has nothing to disclose. C. Jacobs reports grants from MeVis Medical Solutions AG, Bremen, Germany, outside the submitted work. N. Lessmann has nothing to disclose. M.J.M. Broeders has nothing to disclose. M. Silva has nothing to disclose. I. Išgum has nothing to disclose. P.A. de Jong reports departmental research support from Philips Healthcare, during the conduct of the study. M.M. van den Heuvel has nothing to disclose. N. Sverzellati has nothing to disclose. M. Prokop reports personal fees for lectures from Bracco, Bayer, Toshiba and Siemens, grants from Toshiba, other (departmental spin-off with no personal financial interest) from Thiroux, outside the submitted work. U. Pastorino has nothing to disclose. C.M. Schaefer-Prokop has nothing to disclose. B. van Ginneken reports royalties from MeVis Medical Solutions and Delft Imaging Systems, and is co-founder and shareholder of Thirona, outside the submitted work.

### References

- 1 National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011; 365: 395–409.
- 2 Pastorino U, Silva M, Sestini S, *et al.* Prolonged lung cancer screening reduced 10-year mortality in the MILD trial: new confirmation of lung cancer screening efficacy. *Ann Oncol* 2019; 30: 1162–1169.

- 3 de Koning HJ, van der Aalst CM, de Jong PA, *et al.* Reduced lung-cancer mortality with volume CT screening in a randomised trial. *N Engl J Med* 2020; 382: 503–513.
- 4 Cheung LC, Berg CD, Castle PE, *et al.* Life-gained-based versus risk-based selection of smokers for lung cancer screening. *Ann Intern Med* 2019; 171: 623–632.
- 5 Caverly TJ, Cao P, Hayward RA, *et al.* Identifying patients for whom lung cancer screening is preference-sensitive. *Ann Intern Med* 2018; 169: 1–9.
- 6 Oudkerk M, Devaraj A, Vliementhart R, *et al.* European position statement on lung cancer screening. *Lancet Oncol* 2017; 18: e754–e766.
- 7 Moyer VA. Screening for Lung Cancer: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med* 2014; 160: 330–338.
- 8 Harris RP, Sheridan SL, Lewis CL, *et al.* The harms of screening. *JAMA Intern Med* 2014; 174: 281–285.
- 9 Katki HA, Kovalchik SA, Petito LC, *et al.* Implications of nine risk prediction models for selecting ever-smokers for computed tomography lung cancer screening. *Ann Intern Med* 2018; 169: 10–19.
- 10 Katki HA, Kovalchik SA, Berg CD, *et al.* Development and validation of risk models to select ever-smokers for CT lung cancer screening. *JAMA* 2016; 315: 2300–2311.
- 11 Schreuder A, Schaefer-Prokop CM, Scholten ET, *et al.* Lung cancer risk to personalise annual and biennial follow-up computed tomography screening. *Thorax* 2018; 73: 626–633.
- 12 Patz EF, Greco E, Gatsonis C, *et al.* Lung cancer incidence and mortality in National Lung Screening Trial participants who underwent low-dose CT prevalence screening: a retrospective cohort analysis of a randomised, multicentre, diagnostic screening trial. *Lancet Oncol* 2016; 17: 590–599.
- 13 McWilliams A, Tammemagi MC, Mayo JR, *et al.* Probability of cancer in pulmonary nodules detected on first screening CT. *N Engl J Med* 2013; 369: 910–919.
- 14 Raghu VK, Zhao W, Pu J, *et al.* Feasibility of lung cancer prediction from low-dose CT scan and smoking factors using causal models. *Thorax* 2019; 74: 643–649.
- 15 Hassannezhad R, Vahed N. Prediction of the risk of malignancy among detected lung nodules in the National Lung Screening Trial. *J Am Coll Radiol* 2018; 15: 1529–1535.
- 16 Silva M, Milanese G, Sestini S, *et al.* Lung cancer screening by nodule volume in Lung-RADS v1.1: negative baseline CT yields potential for increased screening interval. *Eur Radiol* 2021; 31: 1956–1968.
- 17 Schreuder A, Jacobs C, Lessmann N, *et al.* Combining pulmonary and cardiac computed tomography biomarkers for disease-specific risk modelling in lung cancer screening. *Eur Respir J* 2021; 58: 2003386.
- 18 Gallardo-Estrella L, Pompe E, De Jong PA, *et al.* Normalised emphysema scores on low dose CT: validation as an imaging biomarker for mortality. *PLoS One* 2017; 12: e0188902.
- 19 Lessmann N, van Ginneken B, Zreik M, *et al.* Automatic calcium scoring in low-dose chest CT using deep neural networks with dilated convolutions. *IEEE Trans Med Imaging* 2018; 37: 615–625.
- 20 Charbonnier J-PP, Pompe E, Moore C, *et al.* Airway wall thickening on CT: relation to smoking status and severity of COPD. *Respir Med* 2019; 146: 36–41.
- 21 Lebrecht MB, Balata H, Evison M, *et al.* Analysis of lung cancer risk model (PLCO<sub>M2012</sub> and LLP<sub>v2</sub>) performance in a community-based lung cancer screening programme. *Thorax* 2020, 75: 661–668.
- 22 Robbins HA, Berg CD, Cheung LC, *et al.* Identification of candidates for longer lung cancer screening intervals following a negative low-dose computed tomography result. *J Natl Cancer Inst* 2019; 111: 996–999.
- 23 Gallardo-Estrella L, Lynch DA, Prokop M, *et al.* Normalising computed tomography data reconstructed with different filter kernels: effect on emphysema quantification. *Eur Radiol* 2016; 26: 478–486.
- 24 Mickey RM, Greenland S. The impact of confounder selection criteria on effect estimation. *Am J Epidemiol* 1989; 129: 125–137.
- 25 Zhang Z. Multivariable fractional polynomial method for regression model. *Ann Transl Med* 2016; 4: 174.
- 26 Tammemägi MC, Katki HA, Hocking WG, *et al.* Selection criteria for lung-cancer screening. *N Engl J Med* 2013; 368: 728–736.
- 27 McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 1947; 12: 153–157.
- 28 Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006; 26: 565–574.
- 29 Criqui MH, Knox JB, Denenberg JO, *et al.* Coronary artery calcium volume and density: potential interactions and overall predictive value: the Multi-Ethnic Study of Atherosclerosis. *JACC Cardiovasc Imaging* 2017; 10: 845–854.
- 30 Criqui MH, Denenberg JO, Ix JH, *et al.* Calcium density of coronary artery plaque and risk of incident cardiovascular events. *JAMA* 2014; 311: 271–278.
- 31 Piepoli MF, Hoes AW, Agewall S, *et al.* 2016 European Guidelines on cardiovascular disease prevention in clinical practice: The Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of 10 societies and by invited experts). Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). *Eur Heart J* 2016; 37: 2315–2381.

- 32 Baldwin DR, Duffy SW, Devaraj A, *et al.* Optimum low dose CT screening interval for lung cancer: the answer from NELSON? *Thorax* 2017; 72: 6–7.
- 33 Yousaf-Khan U, van der Aalst C, de Jong PA, *et al.* Final screening round of the NELSON lung cancer screening trial: the effect of a 2.5-year screening interval. *Thorax* 2017; 72: 48–56.
- 34 Kovalchik SA, Tammemagi M, Berg CD, *et al.* Targeting of low-dose CT screening according to the risk of lung-cancer death. *N Engl J Med* 2013; 369: 245–254.
- 35 Balata H, Blandin Knight S, Barber P, *et al.* Targeted lung cancer screening selects individuals at high risk of cardiovascular disease. *Lung Cancer* 2018; 124: 148–153.
- 36 Stevens C, Smith SG, Quaife SL, *et al.* Interest in lifestyle advice at lung cancer screening: determinants and preferences. *Lung Cancer* 2019; 128: 1–5.
- 37 Bugnard F, Ducrot C, Calavas D. Advantages and inconveniences of the Cox model compared with the logistic model: application to a study of risk factors of nursing cow infertility. *Vet Res* 1994; 25: 134–139.
- 38 Moriguchi S, Hayashi Y, Nose Y, *et al.* A comparison of the logistic regression and the Cox proportional hazard models in retrospective studies on the prognosis of patients with gastric cancer. *J Surg Oncol* 1993; 52: 9–13.
- 39 Tammemagi M, Ritchie AJ, Atkar-Khattra S, *et al.* Predicting malignancy risk of screen-detected lung nodules – mean diameter or volume. *J Thorac Oncol* 2019; 14: 203–211.
- 40 Han D, Heuvelmans MA, Oudkerk M. Volume versus diameter assessment of small pulmonary nodules in CT lung cancer screening. *Transl Lung Cancer Res* 2017; 6: 52–61.
- 41 Fan L, Fan K. Lung cancer screening CT-based coronary artery calcification in predicting cardiovascular events. *Medicine* 2018; 97: e10461.
- 42 Budoff MJ, Nasir K, Kinney GL, *et al.* Coronary artery and thoracic calcium on noncontrast thoracic CT scans: comparison of ungated and gated examinations in patients from the COPD Gene cohort. *J Cardiovasc Comput Tomogr* 2011; 5: 113–118.
- 43 American College of Radiology. Lung CT Screening Reporting & Data System v1.1. 2019. Date last accessed: 9 October 2021. [www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Lung-Rads](http://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Lung-Rads)
- 44 Wood DE, Kazerooni EA, Baum SL, *et al.* Lung Cancer Screening, Version 3.2018, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* 2018; 16: 412–441.