Combining pulmonary and cardiac computed tomography biomarkers for disease-specific risk modelling in lung cancer screening

Anton Schreuder[1], Colin Jacobs[1,2], Nikolas Lessmann[1,3], Mireille J.M. Broeders[4,5], Mario Silva[6,7], Ivana Išgum[3], Pim A. de Jong[8], Nicola Sverzellati[7], Mathias Prokop[1,4], Ugo Pastorino[6], Cornelia M. Schaefer-Prokop[1,9] and Bram van Ginneken[1,2,4]

[1]Dept of Radiology, Nuclear Medicine, and Anatomy, Radboud University Medical Center, Nijmegen, The Netherlands. [2]Fraunhofer MEVIS, Bremen, Germany. [3]Image Sciences Institute, University Medical Center Utrecht, Utrecht, The Netherlands. [4]Radboud University Medical Center, Radboud Institute for Health Sciences, Nijmegen, The Netherlands. [5]Dutch Expert Centre for Screening, Nijmegen, The Netherlands. [6]Unit of Thoracic Surgery, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy. [7]Section of Radiology, Unit of Surgical Sciences, Dept of Medicine and Surgery (DiMeC), University of Parma, Parma, Italy. [8]Dept of Radiology, University Medical Center Utrecht, Utrecht, The Netherlands. [9]Dept of Radiology, Meander Medisch Centrum, Amersfoort, The Netherlands.

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

Shareable abstract (@ERSpublications)
Quantitative computed tomography measures of cardiovascular disease and COPD may provide small but reproducible improvements to lung cancer risk prediction accuracy in a screening setting
https://bit.ly/3sWyUMM

Cite this article as: 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 [DOI: 10.1183/13993003.03386-2020].

This article has supplementary material available from erj.ersjournals.com

Received: 7 Sept 2020
Accepted: 18 Jan 2021

Abstract

Objectives Combined assessment of cardiovascular disease (CVD), COPD and lung cancer may improve the effectiveness of lung cancer screening in smokers. The aims were to derive and assess risk models for predicting lung cancer incidence, CVD mortality and COPD mortality by combining quantitative computed tomography (CT) measures from each disease, and to quantify the added predictive benefit of self-reported patient characteristics given the availability of a CT scan.

Methods A survey model (patient characteristics only), CT model (CT information only) and final model (all variables) were derived for each outcome using parsimonious Cox regression on a sample from the National Lung Screening Trial (n=15 000). Validation was performed using Multicentric Italian Lung Detection data (n=2287). Time-dependent measures of model discrimination and calibration are reported.

Results Age, mean lung density, emphysema score, bronchial wall thickness and aorta calcium volume are variables that contributed to all final models. Nodule features were crucial for lung cancer incidence predictions but did not contribute to CVD and COPD mortality prediction. In the derivation cohort, the lung cancer incidence CT model had a 5-year area under the receiver operating characteristic curve of 82.5% (95% CI 80.9–84.0%), significantly inferior to that of the final model (84.0%, 82.6–85.5%). However, the addition of patient characteristics did not improve the lung cancer incidence model performance in the validation cohort (CT model 80.1%, 74.2–86.0%; final model 79.9%, 73.9–85.8%). Similarly, the final CVD mortality model outperformed the other two models in the derivation cohort (survey model 74.9%, 72.7–77.1%; CT model 76.3%, 74.1–78.5%; final model 79.1%, 77.0–81.2%), but not the validation cohort (survey model 74.8%, 62.2–87.5%; CT model 72.1%, 61.1–83.2%; final model 72.2%, 60.4–84.0%). Combining patient characteristics and CT measures provided the largest increase in accuracy for the COPD mortality final model (92.3%, 90.1–94.5%) compared to either other model individually (survey model 87.5%, 84.3–90.6%; CT model 87.9%, 84.8–91.0%), but no external validation was performed due to a very low event frequency.

Conclusions CT measures of CVD and COPD provides small but reproducible improvements to nodule-based lung cancer risk prediction accuracy from 3 years onwards. Self-reported patient characteristics may not be of added predictive value when CT information is available.

Introduction

The three major smoking-related causes of death are cardiovascular disease (CVD), COPD and lung cancer [1, 2]. Validated quantitative computed tomography (QCT) measures of these disease have been described and validated for disease-specific risk prediction [3–7]. Besides external risk factors, lung cancer, CVD and COPD share underlying pathophysiological mechanisms and tend to coexist within similar risk groups [1]. Therefore, a high risk of one of the three diseases may indicate susceptibility to the others.

Lung cancer CT screening reduces the number of lung cancer deaths in high-risk populations where CVD and COPD are important competing causes of death [8–12]. In the National Lung Screening Trial (NLST), 24.1% of all deaths were caused by neoplasms of bronchus and lung, 24.8% by CVD and 10.4% by respiratory illnesses [9]. When a chest CT is performed for detecting and characterising lung nodules, the same scan can be used to extract additional information on QCTs of CVD and COPD.

Pre-scan risk models have been described for predicting lung cancer and competing mortality [13, 14]; their purpose is to select eligible screening participants from the general population. With access to baseline CT findings, post-scan risk stratification can be used to refine previous risk predictions [15]. Several studies have hereby derived lung cancer incidence risk models using baseline CT findings to encourage personalised follow-up interventions [16–21]. To date, no available models have attempted to combine QCTs of lung cancer, CVD and COPD to predict outcomes related to each individual disease. We hypothesised that this additional information would further improve disease-specific risk stratification.

We hereby derived risk models for predicting these outcomes using QCTs of all three diseases. Multiple models for predicting lung cancer incidence, CVD death and/or COPD death were derived to quantify the added value of (post-scan) CT information in addition to (pre-scan) self-reported patient characteristics. A secondary objective was to demonstrate the best performing models' abilities to stratify lung cancer screening participants into groups most and least likely to benefit from disease-specific early interventions.

Methods

Scans and data

NLST data were used to form the derivation cohort (ClinicalTrials.gov NCT00047385) [9]. The NLST was the first randomised controlled trial showing a significant reduction of lung cancer deaths and overall deaths by annual chest CT scans (n=26 722) compared to annual chest radiography (n=26 732) in a high-risk population. There were three annual screening rounds and a subsequent follow-up period of 5 years between August 2002 and December 2009; the median follow-up time was 6.5 years. The use of NLST data for this project was approved for up to 15 000 participants by the National Cancer Institute Cancer Data Access System under project ID NLST-437.

The Multicentric Italian Lung Detection (MILD) trial data were used to externally validate the models (ClinicalTrials.gov NCT02837809) [11, 22–24]. MILD was a randomised controlled trial which followed 4099 participants in Milan. MILD was the first trial to report a significant reduction in lung cancer and overall mortality beyond the fifth year of screening [11]; participants were followed between December 2005 and June 2018. Two intervention groups underwent annual (n=1723) or biennial CT screening (n=1186). The median active screening by volumetric low-dose CT was 6 years; 93.5% of the participants were followed for 9 years.

Model variables

Models were derived to predict three patient outcomes: lung cancer incidence, CVD mortality and COPD mortality. The underlying, immediate or antecedent causes of death were provided, meaning that each participant's death may have been attributed to multiple causes. In MILD, the absolute frequency of COPD deaths was considered insufficient for external validation purposes (n=2 at 5 years' follow-up). Additionally, models for predicting lung cancer mortality were derived and are described in supplementary tables S1–S3, S12–S15.

Variables for model derivation were selected based on the literature and what was available in the datasets [13, 14, 16–21]. These were summarised into four groups: patient characteristics, nodule features, QCTs of CVD and QCTs of COPD (table 1). Patient characteristics and nodule features were obtained from the NLST and MILD datasets; symptoms were not available. If more than one nodule was recorded, the features of the nodule with the longest diameter were used; subjects without a nodule were given a null value for all features. Note that the NLST only reported noncalcified nodules of $\geqslant$4 mm in longest diameter, while MILD reported the diameters and volume for baseline noncalcified nodules of $\geqslant$20 mm^3 by semi-automatic segmentation [22].

TABLE 1 Distribution of variables in the derivation and validation cohorts

	Derivation cohort (NLST)		Validation cohort (MILD)		Intercohort statistics	
	Deceased	Non-deceased	Deceased	Non-deceased	p-value	Effect size[#]
Patients	1598	21 498	154	2133		
Patient characteristics						
Age years	63.6±5.5	61.2±4.9	62.0±6.6	57.2±5.7	<0.001	0.036
Female sex	473 (29.7)	9036 (42.0)	37 (24.0)	688 (32.3)	<0.001	0.055
Race or ethnicity					<0.001	0.100
White	1430 (89.8)	19 247 (89.5)	154 (100)	2129 (99.8)		
Black	89 (5.6)	962 (4.5)	0 (0)	0 (0)		
Asian	20 (1.3)	528 (2.5)	0 (0)	2 (0.1)		
Hispanic	11 (0.7)	331 (1.5)	0 (0)	1 (0.0)		
Mixed or other	42 (2.6)	436 (2.0)	0 (0)	1 (0.0)		
Educational level 0–5[¶]	2.5±1.6	2.7±1.5	1.2±1.4[¶¶]	1.4±1.4[¶¶]	<0.001	0.067
BMI kg·m^{-2}	27.0 (24.0–30.7)	27.3 (24.4–30.5)	25.9 (23.1–28.4)	25.7 (23.5–28.4)	<0.001	0.015
Current smoker	934 (58.7)	10 085 (46.9)	115 (74.7)	1451 (68.0)	<0.001	0.119
Smoking intensity pack-years	56 (44–80)	48 (39–66)	46 (37–66)	39 (31–50)	<0.001	0.024
Smoking duration years	43.5±7.4	39.5±7.3	43.1±7.0	37.9±6.6	<0.001	0.004
Smoking quit time years	7 (3–11)	6 (2–11)	5 (3–7)	5 (3–8)	<0.001	0.034
Lung cancer in family[§]					0.059	<0.001
1	292 (18.3)	3937 (18.3)	32 (20.8)	517 (24.2)		
⩾2	71 (4.5)	671 (3.1)	NA	NA		
Work asbestos	119 (7.5)	956 (4.4)	11 (7.1)	181 (8.5)	<0.001	0.049
COPD diagnosis[f]	433 (27.2)	3732 (17.4)	41 (26.6)	248 (11.6)	<0.001	0.040
Asthma diagnosis	174 (10.9)	2086 (9.7)	16 (10.4)	138 (6.5)	<0.001	0.033
Diabetes diagnosis	256 (16.1)	1951 (9.1)	20 (13.0)	111 (5.2)	<0.001	0.038
Heart disease diagnosis	342 (21.5)	3560 (16.6)	36 (23.4)	244 (11.4)	<0.001	0.003
Hypertension diagnosis	675 (42.4)	7353 (34.2)	60 (39.0)	565 (26.5)	<0.001	0.045
Stroke diagnosis	98 (6.2)	525 (2.4)	8 (5.2)	14 (0.7)	<0.001	0.032
Nodule CT features						
Nodule attenuation					<0.001	0.186
No nodule	1106 (69.2)	15 790 (73.4)	65 (42.2)	933 (43.7)		
Solid	394 (24.6)	4394 (20.4)	65 (42.2)	950 (44.5)		
Part-solid	30 (1.9)	343 (1.6)	5 (3.2)	56 (2.6)		
Non-solid	68 (4.1)	971 (4.5)	19 (12.3)	194 (9.1)		
Longest diameter mm[+]	8 (6–13)	6 (5–9)	6.7 (3.9–11.1)	4.9 (3.1–7.4)	<0.001	0.006
Perpendicular diameter mm[+]	6 (4–10)	5 (4–7)	5.0 (2.8–7.7)	3.9 (2.8–5.6)	<0.001	0.007
Nodule in upper lobe[+]	233 (44.7)	3351 (38.7)	49 (61.3)	474 (39.5)	<0.001	0.105
Nodule spiculation[+]	111 (21.3)	708 (11.7)	NA	NA	NA	NA
Nodule count[+]	1 (1–2)	1 (1–2)	1 (1–2)	1 (1–2)	<0.001	0.006
Quantitative CT measures of CVD						
Coronary calcium volume mm^3	184 (30–681)	45 (0–259)	109 (8–505)	22 (0–141)	<0.001	0.004
Coronary mean calcium density HU	213 (226–266)	206 (0–250)	279 (205–340)	254 (0–309)	<0.001	0.006
Transthoracic aorta calcium volume mm^3	1089 (305–3002)	389 (85–1238)	900 (182–2646)	190 (43–642)	<0.001	0.004
Transthoracic aorta mean calcium density HU	324 (274–376)	311 (250–378)	436 (388–508)	434 (361–523)	<0.001	0.035
Mitral valve calcium volume mm^3	0 (0–11)	0 (0–0)	0 (0–2)	0 (0–0)	<0.001	0.004
Mitral valve mean calcium density HU	0 (0–190)	0 (0–0)	0 (0–216)	0 (0–0)	<0.001	0.001
Aortic valve calcium volume mm^3	0 (0–15)	0 (0–0)	0 (0–14)	0 (0–0)	<0.001	<0.001
Aortic valve mean calcium density HU	0 (0–177)	0 (0–0)	0 (0–233)	0 (0–0)	0.014	<0.001
Quantitative CT measures of COPD						
Total lung volume L	5.68 (4.69–6.77)	5.38 (4.52–6.40)	5.98 (5.18–6.84)	5.92 (5.08–6.83)	<0.001	0.011
Mean lung density[##] HU	−836 (−858–−810)	−839 (−858–−815)	−846 (−862–−825)	−846 (−861–−828)	<0.001	0.007
Emphysema score	0.5 (0.1–2.9)	0.2 (0.1–1.1)	0.1 (0.0–0.5)	0.0 (0.0–0.2)	<0.001	0.047
Pi10	3.0 (2.5–3.5)	2.7 (2.3–3.3)	2.5 (2.3–2.9)	2.4 (2.1–2.6)	<0.001	0.056

Data are presented as n, mean±SD, n (%) or median (interquartile range), unless otherwise stated. NLST: National Lung Screening Trial; MILD: Multicentric Italian Lung Detection; BMI: body mass index; CT: computed tomography; CVD: cardiovascular disease; Pi10: measure of bronchial wall thickness; NA: not applicable. [#]: t-test and coefficient of determination (r^2) for continuous variables with a normal distribution; Mann–Whitney U-test and r^2 for continuous variables with a non-normal distribution; Pearson's Chi-squared test and Cramér's V for categorical variables; [¶]: 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; [+]: of those applicable; regarding nodule features, applies to only the nodule with the longest diameter; [§]: number of first-degree family members diagnosed with lung cancer (a value of "2" was given when two or more family members were diagnosed); [f]: includes prior diagnosis of COPD, emphysema and/or chronic bronchitis; [##]: centred at −1000 HU; [¶¶]: on a scale of 0–4.

QCTs of CVD and COPD were extracted from the CT images automatically using previously described methods [3, 25, 26] (table 1). Calcium volume and mean density were obtained for the coronary arteries (combined), mitral valve, aortic valve and transthoracic aorta [26]. Emphysema score was defined as the percentage of lung voxels below −950 HU after resampling the CT images to 3 mm slice thickness, normalisation and bullae analysis [25]. Airway wall thickness at an internal perimeter of 10 mm (Pi10) was computed as a measure of bronchial wall thickness, as follows: 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 [3].

Dataset formation

The NLST subject inclusion criterion was the availability of a baseline CT image of slice thickness ⩽2.5 mm. Participants with missing data on lung cancer incidence, vital status, time of event, QCTs of CVD or QCTs of COPD were excluded from the study (supplementary figure S1). Cases with unrealistic QCT values were assumed to be algorithm failures and therefore excluded, *i.e.* mean lung density >−300 HU (n=8), mean lung density <−1000 HU (n=3), Pi10 <0.8 (n=190) and Pi10 >6.5 (n=137) [3, 27, 28]. Five-fold multiple imputations were performed to deal with other missing data.

Of those still eligible, all participants who were diagnosed with lung cancer, all who died within the study period and a random sample of all other participants from the CT screening arm up to the maximum number allowed for a single study (15 000 unique subjects) were included in the NLST cohort. The participants sampled at random (alive and lung cancer free) were sampled without replacement and added to the 15 000 unique subjects to simulate the full NLST cohort with a baseline scan (n=23 096). This was to maintain the original probabilities of events which occurred in the NLST, in turn preventing the models from overestimating the risk. All 2287 eligible participants from MILD were used for validation; 2271 (99.3%) out of the 2287 scans had 1 mm slice thickness. Unlike with the derivation cohort, missing or outlier QCTs were replaced with the median values from the validation cohort; other missing data were imputed.

Statistical analysis

Statistical analysis was performed in R (version 3.4.3). Cox proportional hazards regression was performed to derive the models. As each participant could have multiple causes of death, competing risks were not considered in our analysis. The level of significance for including variables in the model was set at α_1=0.20; first-degree fractional polynomials were considered for continuous variables (α_2=0.05) [29]. Three parsimonious models were derived for each outcome: a "survey model" (self-reported patient characteristics only); a "CT model" (nodule features, QCTs of CVD and COPD, age and sex); and a "final model" (all variables). An additional "nodule model" was derived for predicting lung cancer incidence (patient characteristics and nodule features). The proportional hazards assumption for each model was tested by correlating Schoenfeld residuals with time [30]. A graphical diagnostic assessment was performed in cases where there were indications for dependence between residuals and time ($p<0.05$).

β-Coefficients, hazard ratios, estimated baseline hazard functions and calibrated equations to estimate survival probabilities are reported in the supplementary material ("Risk model equations" section). For the purpose of testing accuracy, Kaplan–Meier curves were plotted for each model; each was divided into low-risk (quantiles <0.5), medium-risk (quantiles 0.5–0.9) and high-risk (quantiles >0.9) groups. Sensitivity, specificity, positive predictive value and negative predictive value were calculated at the 0, 0.1, 0.25, 0.5, 0.75 and 0.9 quantile cut-off points at 1, 2, 3 and 5 years' follow-up time (supplementary tables S1 and S2). The time-dependent area under the receiver operating characteristics curve (AUC) and pointwise 95% confidence intervals were calculated; statistical comparison of AUCs were performed using the bootstrap test for paired samples (500 times). Calibration was tested using calibration plots, where the estimated survival probabilities were plotted against the actual survival probabilities. Survival time decision curve analyses are included in the supplementary material (figures S7 and S8) [31].

Comparisons of the variables' distributions between the derivation and validation cohorts was performed using the t-test (normally distributed continuous variables), Mann–Whitney U-test (non-normally distributed continuous variables) or Pearson's Chi-squared test (categorical variables). Effect sizes (ES) were reported in the form of the adjusted coefficient of determination (r^2) for continuous variables (small 0.01<ES<0.09; medium 0.09<ES<0.25; large ES>0.25) and Cramér's V (φ_c) for categorical variables (small 0.1<ES<0.3; medium 0.3<ES<0.5; large ES>0.5) [32, 33]. Nodule spiculation (yes *versus* no) was not available in the validation cohort and was given a value of 0; the number of first-degree family members with lung cancer was binarised (0 *versus* ⩾1).

The same tests used to assess accuracy and calibration in the derivation cohort were used in the validation cohort. The mean and 90th quantile of absolute errors were measured as additional measures of calibration. Internal validation in the form of AUC bootstrap was performed for the CVD and COPD mortality models to assess optimism. Linear regression was performed to assess trends between risk scores.

The discriminatory performance of our lung cancer incidence and CVD mortality models were compared to the Brock nodule risk model [34] and the CVD event risk model by Mets *et al.* [35], respectively. Due to a lack of clinical and spirometry data, existing tools for predicting COPD mortality could not be calculated for comparison. Comparisons between two models were considered statistically significant when a p-value <0.05 was calculated.

More details on the methods can be found in the supplementary material.

Results

Cohort demographics

We formed a derivation cohort of 23 096 resampled participants (supplementary figure S1), where 923 (4.0%) were diagnosed with lung cancer, 392 (1.7%) were lung cancer deaths, 635 (2.7%) CVD deaths, 177 (0.8%) COPD deaths, and 518 (2.2%) deaths from other causes. 10.9% of the participants who died (118 out of 1080) had multiple causes of death (lung cancer, CVD and/or COPD). The validation cohort consisted of 2287 participants, where 108 (4.7%) were diagnosed and 48 (2.1%) died of lung cancer, 54 (2.4%) died of CVD, and 13 (0.6%) died of COPD. Descriptive statistics within and between cohorts are summarised in table 1. The distribution of all but one variable (97.3%; 36 out of 37) between the derivation and validation cohorts were significantly different ($p<0.05$) due to different selection criteria, but the effect sizes were small (<0.09) in 92% (34 out of 37) of the variables [32]. This indicates that the observed effect may be due to a large sample size and should not be overvalued.

Table 2 reports the models' time-dependent AUCs; supplementary table S7 summarises the variables included in each model; and more model details can be found in supplementary tables S8–S21. Other measures of accuracy are reported in supplementary tables S1 (derivation cohort) and S2 (validation cohort). Models were abbreviated based on disease (lung cancer, CVD or COPD), event type (incidence (i) or mortality (m)) and variables considered (survey, CT or final). Statistical testing suggested that nodule features violated the proportional hazards assumption, but graphical assessment only demonstrated a weak pattern with time and was not considered to be an issue (supplementary figure S2).

Lung cancer incidence

Information extracted from CT scans significantly outperformed self-reported characteristics for the stratification of lung cancer incidence (LCi) in the derivation cohort, but the difference in discriminative performance decreased with time (table 2 and figure 1, supplementary tables S1 and S2 and figure S4): The AUC (95% CI) of LCi_{CT} dropped from 93.1% (91.4–94.9%) to 82.5% (80.9–84.0%) between 1 and 5 years' follow-up, whereas that of LCi_{survey} remained stable at 69.6% (66.4–72.8%) and 70.6% (68.8–72.4%), respectively. The AUC of LCi_{final} was significantly greater than that of LCi_{CT} from 3 years' follow-up onwards. LCi_{nodule} performed equivalently to LCi_{CT} at all time points. The full Brock model with spiculation (calibrated to the NLST by Winter *et al.* [32]) was significantly inferior to LCi_{nodule} and LCi_{final} at all time points [33].

External validation of the LCi models did not result in a significant difference between LCi_{final} and LCi_{CT}, although LCi_{nodule} was inferior to both at 3 and 5 years' follow-up (table 2, supplementary figure S3). The full Brock model without spiculation was significantly inferior to LCi_{nodule}, LCi_{CT} and LCi_{final} at all time points [33]. External calibration of NLST-calibrated LCi models in the MILD cohort revealed that the mean absolute error across the models was no greater than 0.004 (0.9 quantile=0.004–0.008) (supplementary table S5). NLST- and MILD-calibrated calibration plots are displayed in supplementary figures S9 and S10, respectively.

Participants with a higher LCi_{final} risk very often also had a higher LCm_{final} risk (NLST cohort r^2=0.90, 95% CI 0.90–0.91); MILD cohort r^2=0.88, 0.86–0.89) (supplementary table S6). Similar discriminative performance trends were seen between the models for predicting lung cancer mortality (supplementary table S3 and figures S3 and S6).

CVD and COPD mortality

CT models for CVD and COPD mortality risk prediction had an equivalent discriminative performance to the survey models in the derivation cohort (table 2 and figures 2 and 3, supplementary table S1): at 1-year

TABLE 2 Models' accuracy in the National Lung Screening Trial (NLST) and Multicentric Italian Lung Detection (MILD) cohorts

	1-year AUC %	p-value	3-year AUC %	p-value	5-year AUC %	p-value
NLST cohort (derivation)						
LCi_{survey}	69.6 (66.4–72.8)	<0.001	69.8 (67.7–71.9)	<0.001	70.6 (68.8–72.4)	<0.001
LCi_{CT}	93.1 (91.4–94.9)	0.170	84.8 (83.1–86.5)	0.031	82.5 (80.9–84.0)	<0.001
LCi_{final}	93.7 (92.0–95.4)	Ref.	85.9 (84.3–87.5)	Ref.	84.0 (82.6–85.5)	Ref.
LCi_{nodule}	93.3 (91.4–95.1)	0.276	84.5 (82.7–86.3)	<0.001	82.1 (80.4–83.7)	<0.001
Brock 2b [34][#,¶]	92.1 (89.9–94.4)	0.069	80.3 (78.1–82.5)	<0.001	77.0 (75.0–78.9)	<0.001
$CVDm_{survey}$	74.4 (68.5–80.3)	0.366	74.5 (71.6–77.4)	<0.001	74.9 (72.7–77.1)	<0.001
$CVDm_{CT}$	73.6 (67.9–79.3)	0.034	77.1 (74.1–80.1)	0.002	76.3 (74.1–78.5)	<0.001
$CVDm_{final}$	76.4 (70.9–82.0)	Ref.	79.5 (76.6–82.3)	Ref.	79.1 (77.0–81.2)	Ref.
METS *et al.* [35]	69.7 (63.4–75.9)	0.005	72.9 (69.6–76.1)	<0.001	72.6 (70.1–75.0)	<0.001
$COPDm_{survey}$	86.6 (72.0–101.2)	0.410	88.3 (84.0–92.6)	0.028	87.5 (84.3–90.6)	<0.001
$COPDm_{CT}$	83.6 (73.2–94.1)	0.049	87.1 (82.7–91.5)	0.002	87.9 (84.8–91.0)	<0.001
$COPDm_{final}$	89.6 (80.5–98.6)	Ref.	91.3 (87.9–94.7)	Ref.	92.3 (90.1–94.5)	Ref.
MILD cohort (validation)						
LCi_{survey}	75.2 (61.3–89.0)	0.008	75.8 (68.0–83.5)	0.167	74.7 (68.0–81.4)	0.103
LCi_{CT}	89.7 (81.5–97.9)	0.621	79.9 (72.5–87.2)	0.575	80.1 (74.2–86.0)	0.909
LCi_{final}	90.2 (81.0–99.5)	Ref.	80.8 (73.7–87.9)	Ref.	79.9 (73.9–85.8)	Ref.
LCi_{nodule}	89.2 (80.1–98.4)	0.200	79.1 (71.6–86.5)	0.014	77.5 (71.3–83.8)	0.005
Brock 1b [34][+]	81.5 (70.4–92.5)	0.022	68.1 (58.9–77.2)	<0.001	67.6 (60.2–75.0)	<0.001
$CVDm_{survey}$	NA	NA	76.1 (60.2–92.1)	0.902	74.8 (62.2–87.5)	0.619
$CVDm_{CT}$	NA	NA	83.8 (68.3–99.4)	0.034	72.1 (61.1–83.2)	0.992
$CVDm_{final}$	NA	NA	76.8 (56.9–96.7)	Ref.	72.2 (60.4–84.0)	Ref.
METS *et al.* [35][§]	NA	NA	65.9 (34.7–97.2)	0.078	63.7 (49.1–78.4)	0.163

Data are presented as receiver operating characteristic area under the curve (AUC), unless otherwise stated. Pointwise 95% AUC confidence intervals are reported in brackets; p-values were obtained from the bootstrap test for paired samples compared to the final model's AUC for the same outcome and time point. LCi_{survey}: lung cancer incidence survey model; LCi_{CT}: lung cancer incidence computed tomography (CT) model; LCi_{final}: final lung cancer incidence model; LCi_{nodule}: lung cancer incidence nodule model; $CVDm_{survey}$: CVD mortality survey model; $CVDm_{CT}$: CVD mortality CT model; $CVDm_{final}$: final CVD mortality model; $COPDm_{survey}$: COPD mortality survey model; $COPDm_{CT}$: COPD mortality CT model; $COPDm_{final}$: final COPD mortality model; ref.: reference value; NA: not applicable. [#]: full lung cancer incidence model with spiculation [34]; [¶]: calibrated by WINTER *et al.* [32]; [+]: full lung cancer incidence model without spiculation [34]; [§]: 3-year CVD risk model [35].

follow-up, $CVDm_{CT}$ (73.6%, 95% CI 67.9–79.3%) and $COPDm_{CT}$ (83.6%, 73.2–94.1%) had (not significantly) lower AUCs than $CVDm_{survey}$ (74.4%, 68.5–80.3%) and $COPDm_{survey}$ (86.6%, 72.0–100%), respectively. For both outcomes, the final models were significantly superior to the survey and CT models in the third and fifth years of follow-up. 3 years after the baseline scan, the model for 3-year risk of CVD event by METS *et al.* [35] was statistically inferior to both $CVDm_{CT}$ and $CVDm_{final}$.

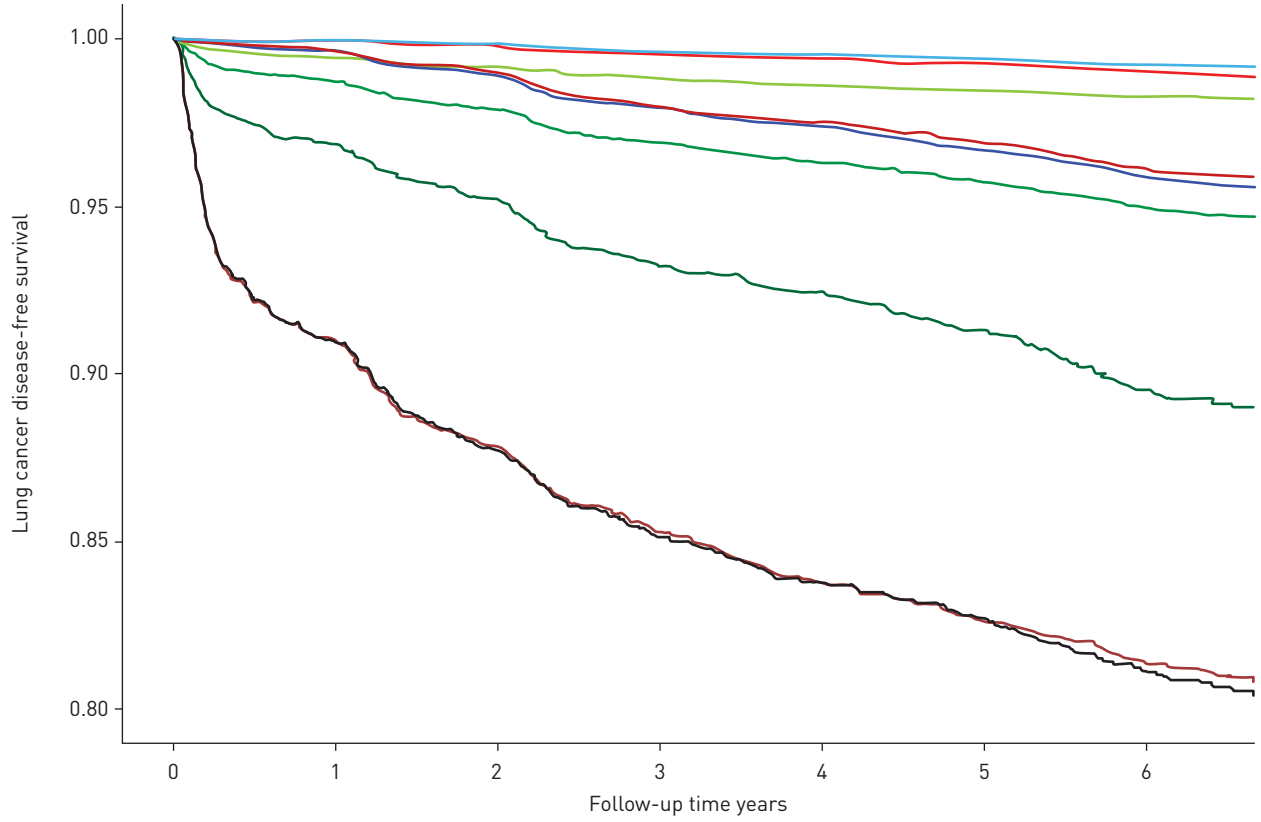
At 3 years' follow-up in the validation cohort, the $CVDm_{CT}$ AUC was significantly higher than that of $CVDm_{survey}$, $CVDm_{final}$, and the model by METS *et al.* [35]. In addition, $CVDm_{CT}$ had the lowest mean absolute error for externally predicted probabilities (0.001; 0.9 quantile=0.002) (supplementary table S5).

To compensate for a lack of external validation, internal validation of the COPD mortality models showed an optimism of no greater than 0.001 at 1-year follow-up and 0.006 at 5 years' follow-up; this indicates a lack of overfitting (supplementary table S4).

Event probability estimates

Among half of the participants in the derivation cohort with the lowest LCi_{final}, $CVDm_{final}$ and $COPDm_{final}$ risks, the average 5-year probabilities of each outcome occurring were 0.6% (61 out of 11 181), 0.5% (54 out of 11 272) and 0.02% (two out of 11 166), respectively (supplementary table S1). Among the 10% with the highest risks, the probabilities were 18.4% (393 out of 2136), 8.2% (172 out of 2102) and 4.0% (82 out of 2033), respectively.

In the validation cohort, half of the validation cohort with the lowest LCi_{final} risk had a 5-year LCi of 0.8% (nine out of 1130); for the half with the lowest $CVDm_{final}$ risk, the CVD mortality probability was 0.4% (four out of 1139). The respective event probabilities among the 10% of the participants with the highest



Number at risk (number of events)

LCi_{survey} <0.5	11 547 (0)	11 429 (61)	11 336 (96)	11 216 (134)	11 100 (155)	10 969 (174)	10 027 (193)
LCi_{survey} 0.5–0.9	9239 (0)	9059 (115)	8899 (195)	8714 (284)	8559 (335)	8352 (387)	7549 (454)
LCi_{survey} >0.9	2310 (0)	2210 (73)	2138 (110)	2044 (153)	1968 (172)	1882 (195)	1679 (230)
LCi_{CT} <0.5	11 548 (0)	11 471 (7)	11 402 (27)	11 295 (52)	11 210 (66)	11 075 (87)	10 141 (113)
LCi_{CT} 0.5–0.9	9238 (0)	9146 (33)	8990 (93)	8800 (183)	8611 (226)	8379 (275)	7552 (343)
LCi_{CT} >0.9	2310 (0)	2081 (209)	1981 (281)	1879 (336)	1806 (370)	1749 (394)	1562 (421)
LCi_{final} <0.5	11 547 (0)	11 481 (6)	11 426 (16)	11 327 (45)	11237 (53)	11 117 (64)	10 141 (84)
LCi_{final} 0.5–0.9	9238 (0)	9138 (34)	8965 (103)	8766 (186)	8585 (239)	8343 (299)	7553 (367)
LCi_{final} >0.9	2311 (0)	2079 (209)	1982 (282)	1881 (340)	1805 (370)	1743 (393)	1561 (426)

FIGURE 1 Lung cancer disease-free survival Kaplan–Meier curves in the derivation cohort (National Lung Screening Trial). Graph and risk table showing the performance of the lung cancer incidence models. The risk from each model is stratified into three risk quantiles divided at the 50th and 90th percentiles. LCi_{survey}: lung cancer incidence survey model; LCi_{CT}: lung cancer incidence computed tomography model; LCi_{final}: final lung cancer incidence model.

model risk probabilities were 12.9% and 3.2% (supplementary table S2). Figure 4 visualises the event probability estimates (calibrated to the relevant cohort) across percentiles.

Discussion

Overall performance of models

We developed and validated risk models for predicting LCi, CVD mortality and COPD mortality. These models are the first to combine objective QCTs of lung cancer, CVD and COPD, namely a consistent backbone of parameters beyond variability of self-reporting [13–16, 18, 19, 36]. The added value of the CT predictors to self-reported patient characteristics was quantified by comparing performances of models without one or the other group of variables (table 2, figure 1). The discriminative ability of each model at

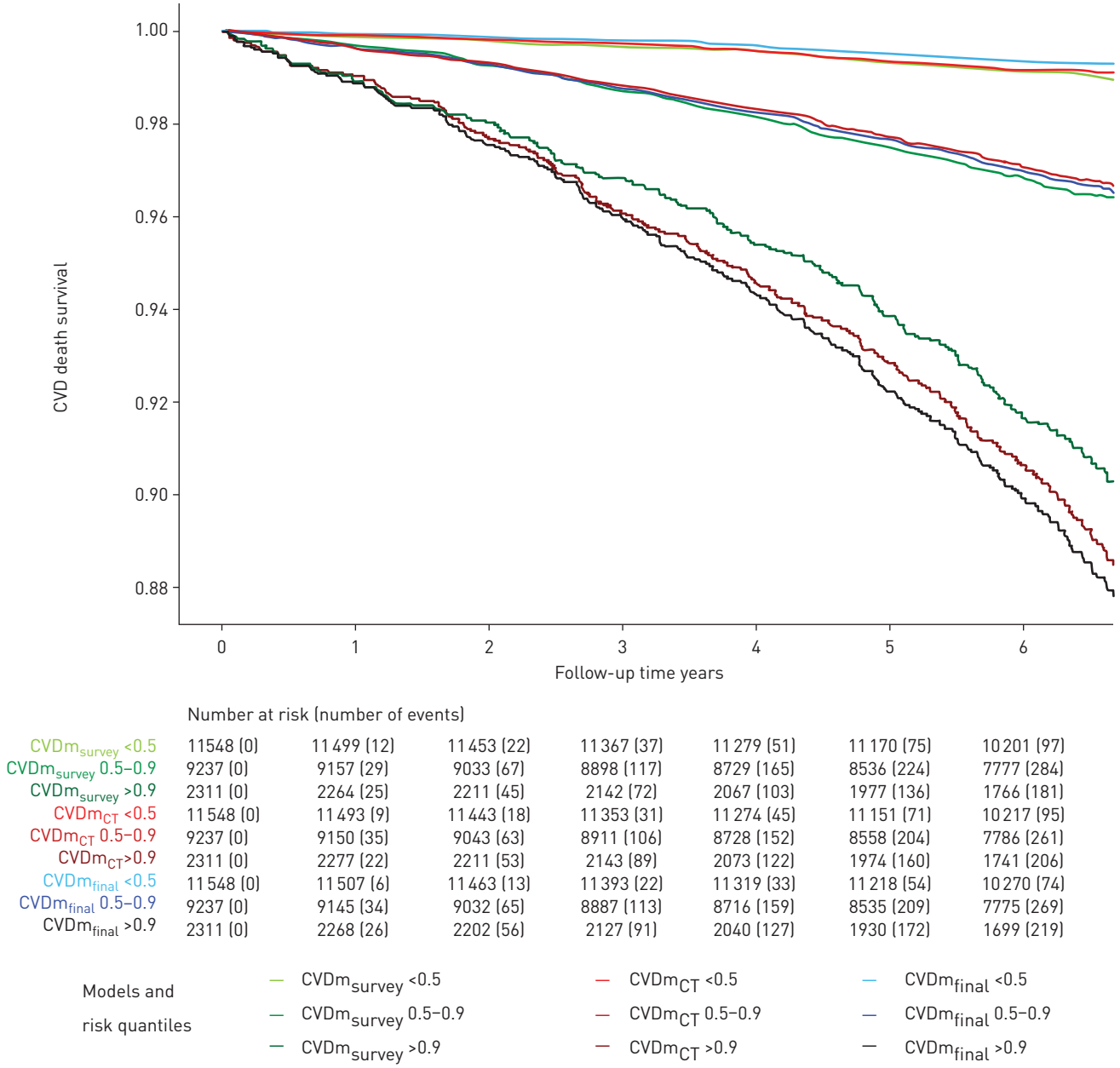
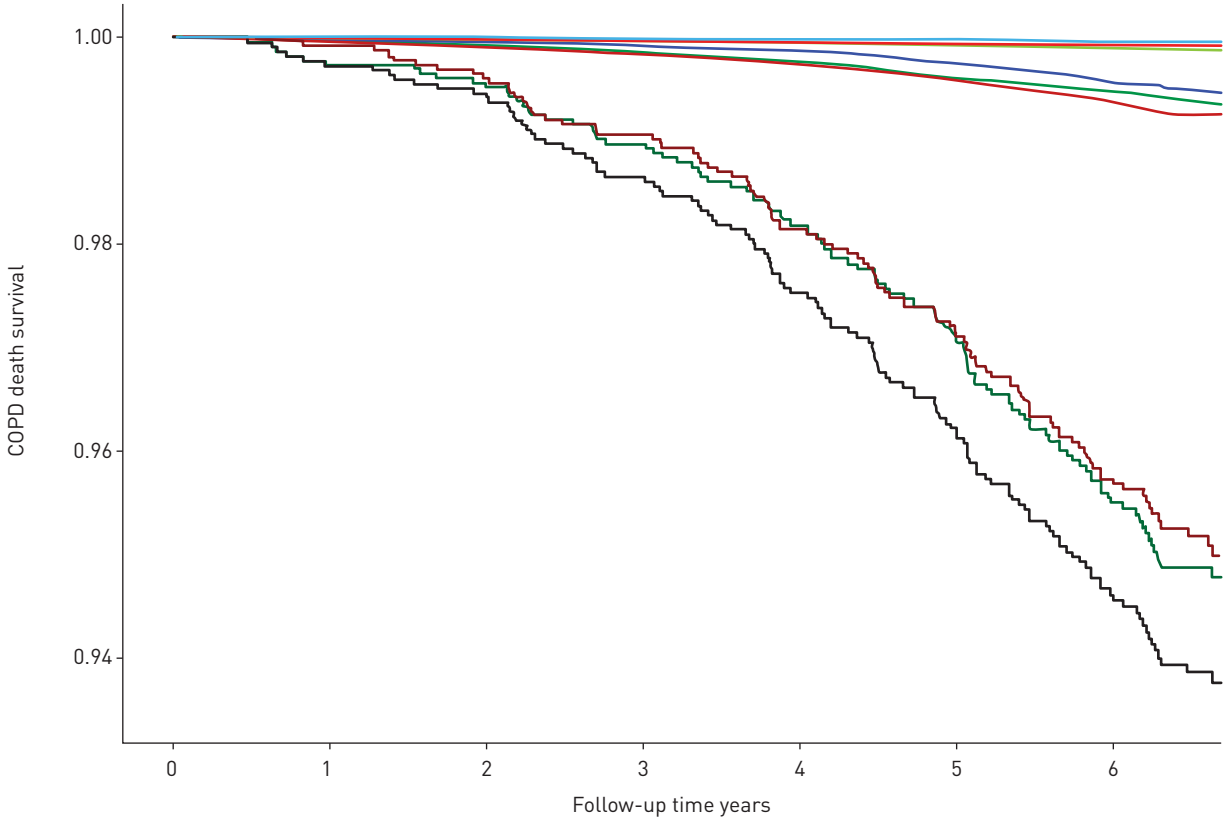


FIGURE 2 Cardiovascular disease (CVD) death survival Kaplan–Meier curves in the derivation cohort (National Lung Screening Trial). Graph and risk table showing the performance of the CVD mortality models. The risk from each model is stratified into three risk quantiles divided at the 50th and 90th percentiles. $CVDm_{survey}$: CVD mortality survey model; $CVDm_{CT}$: CVD mortality computed tomography model; $CVDm_{final}$: final CVD mortality model.

various cut-off points and time points was reported to aid in the selection of thresholds (figures 1–3, supplementary tables S1 and S2 and figure S13).

For predicting events beyond 1 year of follow-up, LCi_{final} was significantly superior to all other LCi models in the derivation cohort (table 2). In the validation cohort, LCi_{CT} performed equivalently to LCi_{final}; both models performed better than LCi_{nodule} from 3 years' follow-up onwards. This suggests that, depending on the cohort, patient characteristics may not improve LCi model discrimination or only contribute to long-term predictions.

For CVD and COPD mortality prediction, the value of combining patient characteristics with CT information in the derivation cohort significantly improved performance compared to either alone at 3 and



Number at risk (number of events)

COPDm$_{survey}$<0.5	11 547 (0)	11 488 (1)	11 421 (2)	11 327 (2)	11 231 (4)	11 104 (7)	10 181 (14)
COPDm$_{survey}$ 0.5-0.9	9239 (0)	9156 (1)	9050 (5)	8915 (14)	8758 (23)	8585 (35)	7788 (46)
COPDm$_{survey}$ >0.9	2310 (0)	2276 (6)	2226 (11)	2165 (23)	2086 (40)	1994 (63)	1775 (95)
COPDm$_{CT}$<0.5	11 547 (0)	11 475 (1)	11 418 (1)	11 322 (2)	11 230 (3)	11 099 (6)	10 169 (7)
COPDm$_{CT}$ 0.5-0.9	9239 (0)	9161 (5)	9048 (8)	8910 (16)	8744 (23)	8556 (37)	7788 (56)
COPDm$_{CT}$ >0.9	2310 (0)	2284 (2)	2231 (9)	2175 (21)	2101 (41)	2028 (62)	1787 (92)
COPDm$_{final}$ <0.5	11 548 (0)	11 503 (0)	11 456 (0)	11 372 (1)	11 287 (1)	11 164 (2)	10 220 (4)
COPDm$_{final}$ 0.5-0.9	9238 (0)	9145 (2)	9028 (5)	8893 (8)	8744 (12)	8568 (21)	7802 (37)
COPDm$_{final}$ >0.9	2310 (0)	2272 (6)	2213 (13)	2142 (30)	2044 (54)	1951 (82)	1722 (114)

FIGURE 3 COPD death survival Kaplan–Meier curves in the derivation cohort (National Lung Screening Trial). Graph and risk table showing the performance of the COPD mortality models. The risk from each model is stratified into three risk quantiles divided at the 50th and 90th percentiles. COPDm$_{survey}$: COPD mortality survey model; COPDm$_{CT}$: COPD mortality computed tomography model; COPDm$_{final}$: final COPD mortality model.

5 years' follow-up (table 2). In the validation cohort, CVDm$_{CT}$ performed best at year 3 of follow-up, but there was no significant differences between the CVD mortality risk models at year 5. Although the combination of patient characteristics and QCTs appeared more beneficial for COPD mortality prediction (supplementary table S4), the lack of external validation is a limitation preventing claims on reproducibility.

External models

We compared our LCi models to the Brock model [34] and our CVD mortality models to the model by Mets *et al.* [35]; our CT and final models consistently outperformed the relevant external models.

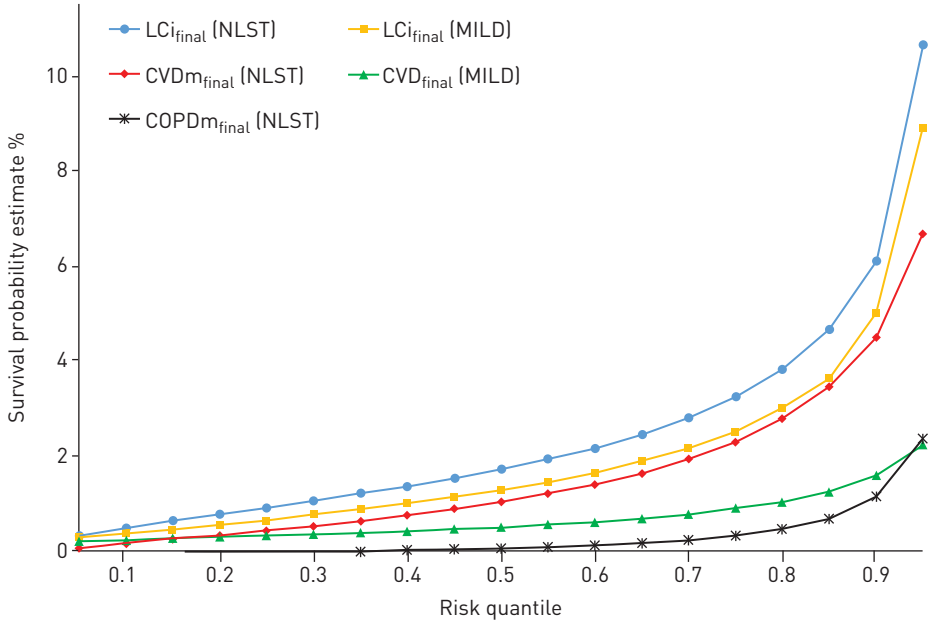


FIGURE 4 Line plots of the calibrated 5-year event probability estimates for all final models. The highest event probabilities were as follows. Final lung cancer incidence model (LCi$_{final}$) (National Lung Screening Trial (NLST)) 60.2%, final lung cancer mortality model (LCm$_{final}$) (NLST) 58.6%, final cardiovascular disease mortality model (CVDm$_{final}$) (NLST) 56.0%, final COPD mortality model (COPDm$_{final}$) (NLST) 26.0%, LCi$_{final}$ (Multicentric Italian Lung Detection (MILD)) 54.7%, LCm$_{final}$ (MILD) 24.9% and CVDm$_{final}$ (MILD) 9.6%. Models were derived in the NLST cohort and validated in the MILD cohort. Models were calibrated to 5 years' follow-up in their respective cohorts.

Although these external models were deemed to be the most suitable for comparison, we acknowledge that the comparisons were not completely fair: the Brock model provides malignancy risk scores for individual nodules (and is therefore not suited for nodule-free participants) [34], and the model by Mets *et al.* [35] provides 3-year risk probabilities for CVD events (including nonfatal diseases).

Assessment of model predictors

QCTs of lung cancer, CVD and COPD were predictive of LCi, but nodule features did not improve the performance of CVD and COPD mortality outcomes. For LCi risk prediction, nodule features (especially nodule diameter) is the most important group of predictors for the short term (⩽2 years); QCTs of CVD and COPD were significant predictors, but only increased the accuracy by a small amount (5-year difference in AUC LCi$_{final}$ and LCi$_{nodule}$ 0.019) (table 2, figure 1 and supplementary tables S10 and S11). Similar findings were described in the Manchester Lung Health Check pilot [37]. All three final models included age, mean lung density, emphysema score, Pi10 and transthoracic aorta calcium volume as predictors, suggesting that the contribution of inflammatory lung and vascular damage might be linked to a phenotypical pattern which facilitates one of the selected outcomes.

Another finding in our study is that the β-coefficients of aorta calcium mean density and coronary calcium mean density were negative in all applicable models (supplementary tables S9, S10, S17, S18, S20 and S21). This is noteworthy because the Agatston score (coronary calcium volume multiplied by the density factor) is the standard for measuring calcium scores [38]. Two studies found that calcified atherosclerotic plaques with a higher density are an indication of stability and were therefore negatively correlated with the CVD event risk [39, 40]. To the best of our knowledge, our study is the first to support this claim in nongated CT images. Note that the calcium mean density of the aortic and mitral valves had positive β-coefficients, which indicates that these biomarkers play a different role in CVD.

An important aspect of this study is that the QCTs of CVD and COPD were obtained fully automatically. Being the only externally calibrated model that did not require human input, the CVDm$_{CT}$ model happened to have the lowest mean absolute error (supplementary table S5). This indicates that the QCTs of CVD and COPD were the most objective and reliable predictors, independent of cohort. We used nodule features

reported by radiologists for this study, but algorithms have already been developed that characterise nodules automatically [41, 42]. The presented models could be useful to select participants with a very high risk for a specific disease and prompt referral to the appropriate specialist.

Clinical relevance

Pulmonary nodule management guidelines recommend additional scans or more invasive tests for detected nodules with a higher malignancy potential [43–45]. Assuming the standard screening interval of 1 year, the added value of our LCi models is not to refine nodule management guidelines, but to downgrade lung cancer risk among screening participants (supplementary figure S13). In practice, participants with a lower lung cancer risk may be recommended a longer interval before their next screening round.

Supplementary figure S13 visualises the costs and benefits if half of the NLST CT cohort with the lowest LCi_{final} risk had been instructed to return for their second screening round two instead of 1 year after baseline [19]. To take it a step further, instructing one-fourth of both the NLST and MILD cohorts with the lowest LCi_{final} risk to return after 5 years would have delayed the lung cancer diagnosis in <4% (25 out of 756 and two out of 59, respectively) of the cases (supplementary tables S1 and S2). Fewer true-negative and false-positive screening tests without significantly reducing the number of screen-detected LCs would improve the overall efficiency of screening.

Besides focusing on lung cancer, the idea of expanding screening to include CVD and COPD is not new [46, 47]. There are secondary prevention drug and lifestyle options with long-term benefits for both CVD and COPD, but studies have demonstrated no added value of screening for early signs of disease [38, 48]. This may be associated with the selection process, where most of the intervention group do not have a sufficiently high *a priori* risk to benefit from the intervention. Our CVD mortality and COPD mortality risk models can be used to refine such a selection procedure, perhaps prompting referral to a specialist. Apart from lifestyle recommendations, the current COPD treatments are symptom dependent [48]. Drugs are generally recommended for patients with a 10-year CVD death risk >5% [38]; the Dutch–Belgian screening (NELSON) trial found that 43–64% of participants with an intermediate-to-high risk of CVD events did not receive antihypertensive drugs or statins [49]. The selection of participants for secondary preventative interventions against CVD and COPD should remain relatively specific (*e.g.* top 10% risk quantile) to avoid overtreatment. If lung cancer is diagnosed, CVD and COPD mortality risk can also be considered when evaluating treatment options such as chemotherapy (cardiotoxicity) and surgery (cardiopulmonary fitness).

Future directions

The present analysis is based on the baseline CT only. From the subsequent screening round onwards, repeat scans enable the quantification of the longitudinal evolution of chest abnormalities. Future models should consider the rate of change between two scans to further improve prediction accuracy. If performed in parallel, outcomes from other screening modalities should also be considered.

Building upon the notion that not all participants will benefit from lung cancer screening, a future study should attempt to identify such a subgroup for re-evaluating their screening eligibility. We performed a shallow analysis where a composite model consisting of LCi_{final}, $CVDm_{final}$ and $COPDm_{final}$ probabilities was used to identify participants with a high risk of lung cancer-free death (supplementary material: "Combining disease-specific risk probabilities: an assessment of supplementary figures S11 and S12"). This group may be identifiable based on a relatively low lung cancer risk and high risk of competing deaths.

Limitations

The main limitation of this study is the low (<100) frequency of each event of interest in the validation cohort. However, besides the NLST, the MILD dataset is larger than other publicly available CT screening cohorts. Furthermore, it is difficult to estimate model performance in a prospective setting because there are many other variables that influenced the outcomes, *e.g.* inclusion criteria, follow-up protocols and all additional tests and interventions performed outside of the screening setting. It is likely that participants' CVD and COPD risk were (externally) assessed at some point during the trial period, which may have affected diagnostic follow-up decisions for some; the contribution of QCTs of CVD and COPD in our lung cancer risk models may therefore be underestimated. The presented risk models have been calibrated to NLST and MILD data, but will probably require retraining or recalibration for other screening programmes.

Although not prone to reader variability, the accuracy of the extracted QCTs is unclear. This especially applies to scans with slice thickness >1 mm. The fact that the CT scans were not ECG-gated reduces the reliability of calcium scores.

Another limitation is that only nodule features from the nodule with the longest diameter were considered. Including the features of multiple nodules may improve predictions, but would increase the number of variables to consider, possibly leading to overfitting. In addition, modelling features which were available in any nodule was attempted, as done in another study [19], but this did not improve the model accuracy. Finally, presence of nodule spiculation was not recorded by the MILD cohort and could therefore not be used to validate the relevant models.

Conclusion

In conclusion, we developed time-dependent risk models for predicting LCi, CVD mortality and COPD mortality using combinations of self-reported patient characteristics and QCTs of lung cancer, CVD and COPD among pre-selected lung cancer screening participants. The added value of using QCTs of CVD and COPD to improve LCi predictions were statistically significant, but may not be clinically relevant within 3 years' follow-up. In turn, the accuracy of the CVD and COPD mortality models were not influenced by nodule features. When CT information is already available for predicting events within the next 2 years, it may not be worthwhile to collect patient characteristics for risk stratification purposes in certain screening populations.

The LCi models offer the possibility to personalise screening with longer intervals up to 5 years among low-risk participants. The CVD mortality and COPD mortality models can be used to refer a select, high-risk group for clinical work-up or to consider treatment contraindications. With the possibility to automate the extraction of imaging features, the human workload can be decreased while optimising personalised recommendations. As expected in a screening setting, most participants should be given high assurance of safety and a relatively small target selection should be watched closely.

Acknowledgements: The authors thank Gabriel Humpire Mamani (Radboud University Medical Center, Nijmegen, the Netherlands), Jean-Paul Charbonnier and Leticia Gallardo-Estrella (Thirona, Nijmegen) 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 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 NCI.

Author contributions: A. Schreuder: conceived and designed the analysis, performed the analysis, wrote the paper. C. Jacobs: conceived the analysis, supervised the analysis, collected the data, contributed data and analysis tools, critically appraised the paper. N. Lessmann: conceived the analysis, collected the data, contributed data and analysis tools, critically appraised the paper. M.J.M. Broeders: conceived the analysis, supervised the analysis, critically appraised the paper. M. Silva: conceived the analysis, collected the data, contributed data, critically appraised the paper. I. Išgum: conceived the analysis, contributed analysis tools, critically appraised the paper. P.A. de Jong: conceived the analysis, contributed analysis tools, critically appraised the paper. N. Sverzellati: conceived the analysis, collected the data, contributed data, critically appraised the paper. M. Prokop: conceived the analysis, critically appraised the paper. U. Pastorino: conceived the analysis, collected the data, contributed data, critically appraised the paper. C.M. Schaefer-Prokop: conceived the analysis, critically appraised the paper. B. van Ginneken: conceived the analysis, supervised the analysis, contributed analysis tools, 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 other (research support to institution) from Philips Healthcare, during the conduct of the study. N. Sverzellati has nothing to disclose. M. Prokop reports personal fees for lectures from Bracco, Bayer, Toshiba and Siemens, grants from Toshiba, other (department spin-off) from Thiroux, outside the submitted work. U. Pastorino has nothing to disclose. C.M. Schaefer-Prokop has nothing to disclose. B. van Ginneken reports other (co-founder/shareholder) from Thirona, grants/royalties from Mevis Medical Solutions and Delft Imaging Systems, outside the submitted work.

References

1 Gerhardsson de Verdier M. The big three concept: a way to tackle the health care crisis? *Proc Am Thorac Soc* 2008; 5: 800–805.
2 World Health Organization. The Top 10 Causes of Death. 2018. www.who.int/en/news-room/fact-sheets/detail/the-top-10-causes-of-death Date last accessed: May 6, 2020. Date last updated: December 9, 2020.
3 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.
4 Fan L, Fan K. Lung cancer screening CT-based coronary artery calcification in predicting cardiovascular events: a systematic review and meta-analysis. *Medicine* 2018; 97: e10461.
5 Gallardo-Estrella L, Pompe E, de Jong PA, *et al.* Normalized emphysema scores on low dose CT: validation as an imaging biomarker for mortality. *PLoS One* 2017; 12: e0188902.
6 Takashima S, Sone S, Li F, *et al.* Small solitary pulmonary nodules (⩽1 cm) detected at population-based CT screening for lung cancer: reliable high-resolution CT features of benign lesions. *AJR Am J Roentgenol* 2003; 180: 955–964.
7 Zulueta JJ, Wisnivesky JP, Henschke CI, *et al.* Emphysema scores predict death from COPD and lung cancer. *Chest* 2012; 141: 1216–1223.
8 de Koning HJ, van der Aalst CM, de Jong PA, *et al.* Reduced lung-cancer mortality with volume CT screening in a randomized trial. *N Engl J Med* 2020; 382: 503–513.
9 National Lung Screening Trial Research Team, Aberle DR, Adams AM, *et al.* Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011; 365: 395–409.
10 Nawa T, Fukui K, Nakayama T, *et al.* A population-based cohort study to evaluate the effectiveness of lung cancer screening using low-dose CT in Hitachi city, Japan. *Jpn J Clin Oncol* 2019; 49: 130–136.
11 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.
12 Becker N, Motsch E, Trotter A, *et al.* Lung cancer mortality reduction by LDCT screening – results from the randomized German LUSI trial. *Int J Cancer* 2020; 146: 1503–1513.
13 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.
14 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.
15 Marcus MW, Duffy SW, Devaraj A, *et al.* Probability of cancer in lung nodules using sequential volumetric screening up to 12 months: the UKLS trial. *Thorax* 2019; 74: 761–767.
16 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.
17 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.
18 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.
19 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.
20 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.
21 Tammemägi MC, ten Haaf K, Toumazis I, *et al.* Development and validation of a multivariable lung cancer risk prediction model that includes low-dose computed tomography screening results: a secondary analysis of data from the National Lung Screening Trial. *JAMA Netw Open* 2019; 2: e190204.
22 Sverzellati N, Silva M, Calareso G, *et al.* Low-dose computed tomography for lung cancer screening: comparison of performance between annual and biennial screen. *Eur Radiol* 2016; 26: 3821–3829.
23 Pastorino U, Rossi M, Rosato V, *et al.* Annual or biennial CT screening *versus* observation in heavy smokers: 5-year results of the MILD trial. *Eur J Cancer Prev* 2012; 21: 308–315.
24 Pastorino U, Bellomi M, Landoni C, *et al.* Early lung-cancer detection with spiral CT and positron emission tomography in heavy smokers: 2-year results. *Lancet* 2003; 362: 593–597.
25 Gallardo-Estrella L, Lynch DA, Prokop M, *et al.* Normalizing computed tomography data reconstructed with different filter kernels: effect on emphysema quantification. *Eur Radiol* 2016; 26: 478–486.
26 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.
27 Takx RAP, Vliegenthart R, Hoesein FAAM, *et al.* Pulmonary function and CT biomarkers as risk factors for cardiovascular events in male lung cancer screening participants: the NELSON study. *Eur Radiol* 2015; 25: 65–71.
28 Cressoni M, Gallazzi E, Chiurazzi C, *et al.* Limits of normality of quantitative thoracic CT analysis. *Crit Care* 2013; 17: R93.

29 Zhang Z. Multivariable fractional polynomial method for regression model. *Ann Transl Med* 2016; 4: 174.
30 Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrika* 1982; 69: 239–241.
31 Vickers AJ, Cronin AM, Elkin EB, *et al.* Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. *BMC Med Inform Decis Mak* 2008; 8: 53.
32 Winter A, Aberle DR, Hsu W. External validation and recalibration of the Brock model to predict probability of cancer in pulmonary nodules using NLST data. *Thorax* 2019; 74: 551–563.
33 Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. *J Exp Psychol Gen* 2012; 141: 2–18.
34 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.
35 Mets OM, Vliegenthart R, Gondrie MJ, *et al.* Lung cancer screening CT-based prediction of cardiovascular events. *JACC Cardiovasc Imaging* 2013; 6: 899–907.
36 Tammemägi MC, Katki HA, Hocking WG, *et al.* Selection criteria for lung-cancer screening. *N Engl J Med* 2013; 368: 728–736.
37 Lebrett MB, Balata H, Evison M, *et al.* Analysis of lung cancer risk model ($PLCO_{M2012}$ and LLP_{v2}) performance in a community-based lung cancer screening programme. *Thorax* 2020; 75: 661–668.
38 Piepoli MF, Hoes AW, Agewall S, *et al.* Linee guida europee 2016 sulla prevenzione delle malattie cardiovascolari nella pratica clinica. Sesta Task Force congiunta della Società Europea di Cardiologia e di altre Società sulla Prevenzione delle Malattie Cardiovascolari nella Pratica Clinica (costituita da rappresentanti di 10 società e da esperti invitati). Redatte con il contributo straordinario dell'Associazione Europea per la Prevenzione e Riabilitazione Cardiovascolare (EACPR). [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)]. *G Ital Cardiol* 2017; 18: 547–612.
39 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.
40 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.
41 Charbonnier J-P, Chung K, Scholten ET, *et al.* Automatic segmentation of the solid core and enclosed vessels in subsolid pulmonary nodules. *Sci Rep* 2018; 8: 646.
42 Ciompi F, Chung K, van Riel SJ, *et al.* Towards automatic pulmonary nodule management in lung cancer screening with deep learning. *Sci Rep* 2017; 7: 46479.
43 American College of Radiology. Lung CT Screening Reporting & Data System (LUNG-RADS) v1.1. 2019. Available from: www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Lung-Rads
44 Callister MEJJ, Baldwin DR, Akram AR, *et al.* British Thoracic Society guidelines for the investigation and management of pulmonary nodules. *Thorax* 2015; 70: Suppl. 2, ii1–ii54.
45 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.
46 Ravenel JG, Nance JW. Coronary artery calcification in lung cancer screening. *Transl Lung Cancer Res* 2018; 7: 361–367.
47 Mets OM, de Jong PA, Prokop M. Computed tomographic screening for lung cancer: an opportunity to evaluate other diseases. *JAMA* 2012; 308: 1433–1434.
48 Singh D, Agusti A, Anzueto A, *et al.* Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease: the GOLD science committee report 2019. *Eur Respir J* 2019; 53: 1900164.
49 Jacobs PC, Gondrie MJA, van der Graaf Y, *et al.* Coronary artery calcium can predict all-cause mortality and cardiovascular events on low-dose CT screening for lung cancer. *AJR Am J Roentgenol* 2012; 198: 505–511.