Supplement

Contents

Methods

Data selection (NLST)

We obtained baseline patient characteristics, nodule features, follow-up outcomes, and chest CT images from National Lung Screening Trial (NLST) participants (1). The data was cleaned (missing data was given blank values) and new patient characteristics variables were created, i.e., body mass index (BMI), time since smoking quit, number of first-degree family members diagnosed with lung cancer, diagnosis of any cancer prior to trial, the follow-up time from randomization to event or date last known alive, and disease-specific causes of death (as determined by the endpoint verification process [EVP] and death certificate). Lung cancer mortality, lung cancer incidence, and other causes of death were provided.

In the NLST data set, the variables race, ethnicity and education were separated into the following categories: Race was divided into White, Black or African-American, Asian, American Indian or Alaskan native, Native Hawaiian or other Pacific Islander, and more than one race; ethnicity was binary: Hispanic or Latino, or neither; educational level completed ranged from 8^{th} grade or less, 9^{th} to 11^{th} grade, high school graduate or General Educational Development, post high school training (excluding college), associate degree or some college, bachelor's degree, graduate school, and other. As most other studies and models did not record race and ethnicity as separate variables, we merged ethnicity with race by allocating all participants who had both a Hispanic or Latino ethnicity and White race into a new race category (Hispanic or Latino). Furthermore, due to the low prevalence (<1%) of the race groups "American Indian or Alaskan native" and "Native Hawaiian or other Pacific Islander" and the ambiguity of the "mixed" race group, the three were merged to form the "mixed or other" group. To make the education variable more applicable to other educational systems and due to their low frequencies, "8^{th} grade or less" and "9^{th} to 11^{th} grade" were merged as "did not complete high school," and "graduate school" and "other" were merged to "graduate school or higher." Furthermore, education is a categorical variable but was applied as a continuous variable similar to in the $PLCO_{m2012}$ (2); "11^{th} grade or less" was set zero, plus one per increase in level up to five.

Age (years) and sex (male vs female) were considered for all models. Detailed patient characteristics (which would normally only be obtained via a survey) were race or ethnicity (White, Black or African-American, Asian, Hispanic or Latino, or mixed or other), educational level (range: 0 to 5), BMI (kg/m^2), smoking status (current vs former), smoking intensity (pack years), smoking duration (years), smoking quit time (years, 0 if current smoker), number of first-degree family members diagnosed with lung cancer (range: 0 to 2 [a value of "2" was given when two or more family members were diagnosed due to a very low prevalence of those with three or more]), exposure to asbestos at work for at least one year (yes vs no), and the prior diagnosis of chronic obstructive pulmonary disease (COPD), asthma (child or adult), diabetes, heart disease, hypertension, and stroke (yes vs no). The prior diagnosis of COPD was also considered positive if the subject had a prior diagnosis of chronic bronchitis and/or emphysema. The NLST listed other diseases in their survey, but these were not available in the MILD cohort and were therefore not considered for modelling.

The list of prospectively detected pulmonary nodules as reported by radiologists who participated in the NLST was used. The following nodule features were considered in our study: longest diameter (mm), longest perpendicular diameter (mm), attenuation (solid [soft tissue], nonsolid [ground glass], or partsolid [mixed]), upper lobe location (yes vs no), spiculation (yes vs no), and nodule count. If more than one nodule was recorded, the features of the nodule with the longest diameter was used; subjects who did not have a nodule were given a null value for all nodule features. Note that the NLST only reported non-calcified nodules of at least 4 mm in longest diameter. Nodules reported to have a longest diameter of 20 mm or greater were visually inspected and corrected if necessary.

Validation data set preparation (MILD)

We obtained baseline patient characteristics, nodule features, follow-up outcomes, and chest CT images from Multicentric Italian Lung Detection (MILD) trial participants. The data was cleaned (missing data was given blank values) and the variables were created and transformed to be interchangeable with that of the NLST. Educational level was available in five levels: elementary school graduate, middle school graduate, high school graduate, university attendee, business school graduate, and bachelor's degree. "elementary school graduate" and "middle

school graduate" were classified as "did not complete high school," "university attendee" was classified as "post high school training (excluding college)," and "business school graduate" was classified as "associate degree or some college." The "first-degree family members diagnosed with lung cancer" variable only mentions whether there was at least one or none; the exact number is not given. All other patient characteristics variables selected from the NLST were available.

We utilized the MILD data set of lung nodules retrospectively detected using computer-aided diagnosis, which does not correspond to the data set of nodules retrospectively detected by computer-aided diagnosis (3,4). Only a maximum of five nodules were reported per scan per time point; nodules only counted in the nodule count if larger than or equal to 20 mm^3 in volume. The same nodule features as from the NLST were considered except for spiculation, which was not available and therefore given a null value when applicable. The longest perpendicular diameter was considered the longest diameter in one of the other two planes which was not the plane which the longest diameter was measured.

Quantitative CT measures

The CT images were used to obtain quantitative CT measures (QCT) of CVD – calcium volume and mean density of the coronary arteries, mitral valve, aortic valve, and transthoracic aorta (5,6) – and chronic obstructive pulmonary disease (COPD) – lung volume, mean lung density, normalized emphysema score (the percentage of lung voxels below -950 HU after resampling the CT images to 3mm slice thickness, normalization, and bullae analysis) (7), and bronchial wall thickness (Pi10, the square root of the airway wall area for a theoretical 10mm lumen perimeter airway derived using the linear regression of the square root of segmented wall area against the lumen perimeter)(8). As quality control, cases with extreme outlier values were excluded, i.e., mean lung density > -300 HU, mean lung density < -1000 HU, Pi10 < 0.8, and Pi10 > 6.5.

Data set formation

The primary 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, death status, time of event, nodule features, QCTs of CVD, or QCTs of COPD were excluded from the NLST cohort (Figure 1). All remaining 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 cohort up to a maximum of 15000 unique subjects were included in the NLST cohort. This was due to the limit of NLST CT images used for one project set by the Cancer Data Access System, project ID "NLST-437. Subsequently, the proportion of non-deceased participants without a lung cancer diagnosed were resampled without replacement to simulate the full NLST cohort. This was calculated by taking the proportion of deceased or participants diagnosed with lung cancer included in our study out of those in the CT arm with a baseline scan, then applying that proportion to the non-deceased participants without a lung cancer diagnosis in the CT arm with a baseline scan to obtain the total number of non-deceased participants without a lung cancer diagnosis who should be included after resampling. This was to maintain the original probabilities of events which occurred in the NLST, in turn preventing the models from overestimating the risk. Resampling was not necessary for the validation cohort as there were no limitations on CT image usage. Furthermore, almost all CT images from MILD were available in 1 mm slice thickness (2271/2287, 99.3%).

Mean lung density was centered to –1000 HU (i.e., 1000 was added to the actual value of the mean lung density) to circumvent modelling issues with negative values.

Multiple imputation using the 'mice' function (R package 'mice', 'cart' method) was performed to impute missing data using classification and regression trees. Of the 15000 NLST subjects included in the NLST data set (before resampling), there was missing data from race (n=13), education (n=21), BMI (n=41), first-degree family history of lung cancer (n=623), exposure to asbestos at work (n=20), and the prior diagnoses of COPD (n=43), asthma (n=15), diabetes (n=14), heart disease (n=33), hypertension (n=16), and stroke (n=14).

2303 subjects were considered for the validation cohort. Of these, 8/2303 (0.3%) were not part of the first screening round and 9/2303 (0.4%) were missing baseline scans. Of the 2287 subjects included, some QCTs of CVD (n=24) and Pi10 values (n=132) could not be extracted from the scans; in contrast to excluding these cases as was done with the

NLST cohort, the missing values were replaced with the corresponding median values from the MILD cohort set to avoid excluding cases.

Equations S1: Final lung cancer incidence model equation (LCi_{final})

$$
\begin{aligned}
&Linear\ predictor \\
&= (female \times 0.12611) + \left(\frac{BMI}{10} \times (-0.19128)\right) + \left(\left(\frac{smoking\ intensity}{100}\right)^{-2} \times (-0.10613)\right) \\
&+ \left(\frac{smoking\ duration}{10} \times 0.25895\right) + \left(\frac{smoking\ quit\ time}{10} \times (-0.30764)\right) \\
&+ (lung\ cancer\ in\ family \times 0.18273) + (\ln(emphysema\ score + 0.1) \times 0.18683) \\
&+ \left(\left(\frac{mean\ lung\ density + 1000}{100}\right)^{-1} \times (-0.74125)\right) + (Pi10 \times 0.09703) \\
&+ \left(\ln\left(\frac{aorta\ calcium\ volume + 0.1}{100} \times 0.17668\right)\right) \\
&+ \left(\left(\frac{aorta\ calcium\ mean\ density + 0.1}{100}\right)^{0.5} \times (-0.57768)\right) \\
&+ \left(\left(\frac{nodule\ longest\ diameter + 1}{10}\right)^{-2} \times (-0.44997)\right) \\
&+ \left(\left(\frac{nodule\ longest\ diameter + 1}{10}\right)^{-2} \times ln\left(\frac{nodule\ longest\ diameter + 1}{10}\right) \times (-0.22010)\right) \\
&+ (\ln(nodule\ perpendicular\ diameter + 1) \times 1.20812) + (nodule\ in\ upper\ lobe \times 0.22769) \\
&+ (spiculated\ nodule \times 0.64707)
\end{aligned}
$$

$$
\begin{aligned}
&NLST\ baseline\ hazard\ function\ estimate = H_0(t)_{lung\ cancer\ incidence} \\
&= 0.0018772 + 0.021448 \times \frac{\mathrm{t}+1}{1000} + (-0.0059062) \times \frac{(\mathrm{t}+1)(\ln(\mathrm{t}+1))}{1000}
\end{aligned}
$$

(where t is the follow-up time in days; for 5 years follow-up, this equals 1826)

$$
\begin{aligned}
&Cumulative\ survival\ probability = S(t) \\
&= \exp\left(-exp(linear\ predictor - \ NLST\ mean\ linear\ predictor\ ) \times H_0(t)_{lung\ cancer\ incidence}\right)
\end{aligned}
$$

(where NLST mean linear predictor = 0.10129)

$$
\begin{aligned}
&Five-year\ lung\ cancer\ incidence\ risk\ calibrated\ to\ the\ National\ Lung\ Screening\ Trial \\
&= S(t) \times 0.60343 + 0.39822
\end{aligned}
$$

$$
\begin{aligned}
&Five-year\ lung\ cancer\ incidence\ risk\ calibrated\ to\ the\ Multicentric\ Italian\ Lung\ Detection\ trial \\
&= S(t) \times 0.85413 + 0.14396
\end{aligned}
$$

Table S1: Self-reported demographics-based competing mortality risk model (CD_{survey})

Variable	Beta coefficient	Odds ratio	95% CI	P value
Model intercept	-7.25928	0.00	0.00 to 0.00	<0.001
Age, per year	0.05829	1.06	1.04 to1.08	<0.001
Sex, female	-0.58787	0.56	0.47 to 0.65	<0.001
White race, reference	N/A	N/A	N/A	N/A
Black race, yes	0.22484	1.25	0.917 to 1.68	0.143
Asian race, yes	-1.41522	0.24	0.10 to 0.50	0.001
Hispanic race, yes	-0.50493	0.60	0.26 to 1.19	0.191
Mixed or other race, yes	0.25358	1.29	0.81 to 1.96	0.261
Body mass index, per kg/m^2*	2.30426	10.02	2.12 to 46.00	0.003
Current smoker, yes	0.36904	1.45	1.22 to 1.71	<0.001
Smoking duration, per year*	-0.06592	0.94	0.90 to 0.97	<0.001
Hypertension diagnosis, yes	0.24762	1.28	1.10 to 1.49	0.001
Diabetes diagnosis, yes	0.59631	1.82	1.48 to 2.21	<0.001
Heart disease diagnosis, yes	0.48109	1.62	1.48 to 2.21	<0.001
Stroke diagnosis, yes	0.65990	1.93	1.43 to 2.58	<0.001
Asthma diagnosis, yes	0.32387	1.38	1.10 to 1.72	0.005
COPD diagnosis, yes	0.37986	1.46	1.23 to 1.73	<0.001

* Transformed as follows: $(\text{body mass index}/10)^{-2}$; $(\text{smoking duration}/100)^{-2}$

To calculate the five-year risk probability of lung cancer incidence, first find the sum of the products of each (transformed) variable and their respective beta coefficient to obtain the linear predictor, then insert the value into the following equation: $\frac{1}{1+e^{-(linear\ predictor)}}$

CI, confidence intervals.

Table S2: CT-based competing mortality risk model (CD_{CT})

Variable	Beta coefficient	Odds ratio	95% CI	P value
Model intercept	-3.73187	0.02	0.01 to 0.08	<0.001
Age, per year	0.03097	1.03	1.02 to 1.05	<0.001
Sex, female	-0.47135	0.62	0.53 to 0.74	<0.001
Emphysema score, per point	0.06217	1.06	1.05 to 1.08	<0.001
Bronchial wall thickness (Pi10), per point	0.20689	1.23	1.12 to 1.35	<0.001
Mean lung density, per HU*	-2.02753	0.13	0.07 to 0.26	<0.001
Aorta calcium volume, per mm^3*	0.21075	1.23	1.17 to 1.30	<0.001
Aorta calcium mean density, per HU*	-1.52263	0.22	0.09 to 0.51	0.001
Coronary calcium volume, per mm^{3*}	0.19567	1.22	1.11 to 1.33	<0.001
Mitral valve calcium volume, per mm^3*	0.25931	1.30	1.06 to 1.56	0.009
Mitral valve calcium mean density, per HU*	0.04676	1.05	1.02 to 1.07	<0.001
Solid nodule (largest nodule), present	-0.22020	0.80	0.67 to 0.96	0.0183

* Transformed as follows: $((\text{mean lung density} + 1000)/100)^{-1}$; ln((aorta calcium volume + 0.1)/1000); (aorta calcium mean density/1000); (coronary calcium volume/1000); (mitral valve calcium volume/1000); ln((mitral valve calcium mean density+0.1)/100).

To calculate the five-year risk probability of lung cancer incidence, first find the sum of the products of each (transformed) variable and their respective beta coefficient to obtain the linear predictor, then insert the value into the following equation: $\frac{1}{1+e^{-(linear\ predictor)}}$

CI, confidence intervals.

Table S3: Internal validation competing death risk models

Model	Original sample AUC (%) (95% confidence interval)	Bootstrap sample AUC (%) (95% confidence interval)	Optimism
	5-year	5-year	5-year
CD_{Survey}	0.704 (0.704 to 0.704)	0.710 (0.709 to 0.710)	0.006
CD_{CT}	0.717 (0.717 to 0.717)	0.720 (0.720 to 0.721)	0.003
CD_{final}	0.740 (0.740 to 0.741)	0.746 (0.746 to 0.747)	0.006

One thousand bootstrap replications were performed for each model and time point. The optimism is the difference between the average bootstrap sample AUC and original sample AUC; a greater value indicates more overfitting.

AUC, receiver operating characteristic area under the curve; CD, competing death.

Table S4: Contingency table of outcomes per risk quintile (derivation cohort)

			Lung cancer incidence risk (LCi_{final})				
			Q1: ≤0.79%	Q2: 0.79%-1.38%	Q3: 1.38-2.18%	Q4: 2.18-3.85%	Q5: >3.85%
Total participant count (n=23096)	Competing death risk (CD_{final})	Q1: ≤1.11%	2578 (12)	1129 (5)	515 (2)	234 (1)	140 (1)
		Q2: 1.11-1.90%	1128 (5)	1467 (7)	1058 (5)	623 (3)	281 (1)
		Q3: 1.90%-2.93%	555 (3)	1071 (5)	1237 (6)	1064 (5)	507 (2)
		Q4: 2.93-4.92%	266 (1)	629 (3)	1124 (5)	1407 (6)	888 (4)
		Q5: >4.92%	92 (0)	322 (1)	686 (3)	1291 (6)	1756 (8)
5-year lung cancer diagnoses (n=756)		Q1: ≤1.11%	4 (1)	9 (1)	4 (1)	6 (1)	23 (3)
		Q2: 1.11-1.90%	0 (0)	8 (1)	9 (1)	11 (1)	48 (6)
		Q3: 1.90%-2.93%	3 (0)	9 (1)	18 (2)	33 (4)	80 (11)
		Q4: 2.93-4.92%	4 (1)	5 (1)	21 (3)	46 (6)	121 (16)
		Q5: >4.92%	1 (0)	3 (0)	9 (1)	40 (5)	241 (32)
5-year competing deaths (n=800)		Q1: ≤1.11%	20 (3)	3 (0)	7 (1)	1 (0)	1 (0)
		Q2: 1.11-1.90%	21 (3)	20 (3)	16 (2)	10 (1)	13 (2)
		Q3: 1.90%-2.93%	10 (1)	23 (3)	29 (4)	25 (3)	10 (1)
		Q4: 2.93-4.92%	6 (1)	17 (2)	42 (5)	58 (7)	43 (5)
		Q5: >4.92%	7 (1)	28 (4)	48 (6)	113 (14)	229 (29)
Competing deaths per lung cancer diagnosis (mean=1.06)		Q1: ≤1.11%	5.00	0.33	1.75	0.17	0.04
		Q2: 1.11-1.90%	21/0	2.50	1.78	0.91	0.27
		Q3: 1.90%-2.93%	3.33	2.56	1.61	0.76	0.13
		Q4: 2.93-4.92%	1.50	3.40	2.00	1.26	0.36
		Q5: >4.92%	7.00	9.33	5.33	2.83	0.95

The cells are shaded grey according to the proportion of outcomes which occurred in that cell, where a darker shade indicates a greater proportion. Brackets indicate the percentage of the total count where applicable. The double lines indicate the suggested separation of screening participants into a group which should continue to be screened (right of the double line) and a group with a relatively low lung cancer risk and high risk of competing death which is unlikely to benefit from screening (left of the double line).

CD_{final}, final competing death model; LCi_{final}, final lung cancer incidence model (9); n, total count; Q, quintile (e.g., Q1 = 0^{th} to 20^{th} percentile).

Table S5: Contingency table of outcomes per risk quintile (validation cohort)

			Lung cancer incidence risk (LCi_{final})				
			Q1: ≤0.79%	Q2: 0.79-1.38%	Q3: 1.38-2.18%	Q4: 2.18-3.85%	Q5: >3.85%
Total participant count (n=23096)	Competing death risk (CD_{final})	Q1: ≤1.11%	441 (19)	231 (10)	110 (5)	48 (2)	34 (1)
		Q2: 1.11-1.90%	163 (7)	171 (7)	126 (6)	108 (5)	68 (3)
		Q3: 1.90-2.93%	59 (3)	63 (3)	88 (4)	83 (4)	69 (3)
		Q4: 2.93-4.92%	38 (2)	33 (1)	46 (2)	78 (3)	80 (3)
		Q5: >4.92%	4 (0)	17 (1)	19 (1)	42 (2)	68 (3)
5-year lung cancer diagnoses (n=756)		Q1: ≤1.11%	1 (2)	0 (0)	0 (0)	0 (0)	1 (2)
		Q2: 1.11-1.90%	1 (2)	3 (5)	3 (5)	3 (5)	8 (14)
		Q3: 1.90-2.93%	0 (0)	2 (3)	1 (2)	1 (2)	2 (3)
		Q4: 2.93-4.92%	0 (0)	2 (3)	0 (0)	3 (5)	11 (19)
		Q5: >4.92%	0 (0)	0 (0)	2 (3)	4 (7)	11 (19)
5-year competing deaths (n=800)		Q1: ≤1.11%	2 (6)	1 (3)	1 (3)	0 (0)	0 (0)
		Q2: 1.11-1.90%	1 (3)	1 (3)	0 (0)	1 (3)	2 (6)
		Q3: 1.90-2.93%	0 (0)	2 (6)	1 (3)	2 (6)	0 (0)
		Q4: 2.93-4.92%	2 (6)	0 (0)	0 (0)	6 (18)	2 (6)
		Q5: >4.92%	1 (3)	1 (3)	3 (9)	1 (3)	3 (9)
Competing deaths per lung cancer diagnosis (mean=1.06)		Q1: ≤1.11%	2.00	1/0	1/0	0/0	0.00
		Q2: 1.11-1.90%	1.00	0.33	0.00	0.33	0.25
		Q3: 1.90-2.93%	0/0	1.00	1.00	2.00	0.00
		Q4: 2.93-4.92%	2/0	0.00	0/0	2.00	0.18
		Q5: >4.92%	1/0	1/0	1.50	0.25	0.27

The cells are shaded grey according to the proportion of outcomes which occurred in that cell, where a darker shade indicates a greater proportion. Brackets indicate the percentage of the total count where applicable. The double lines indicate the suggested separation of screening participants into a group which should continue to be screened (right of the double line) and a group with a relatively low lung cancer risk and high risk of competing death which is unlikely to benefit from screening (left of the double line).

CD_{final}, final competing death model; LCi_{final}, final lung cancer incidence model (9); n, total count; Q, quintile (e.g., Q1 = 0^{th} to 20^{th} percentile).

Table S6: Clinical outcomes of risk stratification

Risk models used	Risk group	Participants (%)	5-year LC diagnoses (%)	5-year CDs (%)	Number needed to screen to detect 1 LC	CDs per LC diagnosis
Derivation cohort (NLST)						
LCi_{final} & CD_{final}	High-LC-low-CD	16838 (73)	727 (96)	643 (80)	23	0.88
	Low-LC-high-CD	6258 (27)	29 (4)	157 (20)	216	5.41
LCi_{final} & CD_{CD}	High-LC-low-CD	16537 (72)	723 (96)	639 (80)	23	0.88
	Low-LC-high-CD	6559 (28)	33 (4)	161 (20)	199	4.90
LCi_{final} & CD_{survey}	High-LC-low-CD	16530 (72)	725 (96)	643 (80)	23	0.89
	Low-LC-high-CD	6566 (28)	31 (4)	157 (20)	212	5.06
LCi_{final} & CDRAT	High-LC-low-CD	16638 (72)	720 (95)	671 (84)	23	0.93
	Low-LC-high-CD	6458 (28)	36 (5)	129 (16)	179	3.58
	Full derivation cohort	*23096 (100)*	*756 (100)*	*800 (100)*	*31*	*1.06*
Validation cohort (MILD)						
LCi_{final} & CD_{final}	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
LCi_{final} & CD_{CD}	High-LC-low-CD	1538 (67)	54 (92)	23 (70)	28	0.43
	Low-LC-high-CD	749 (33)	5 (8)	10 (30)	150	2.00
LCi_{final} & CD_{survey}	High-LC-low-CD	1414 (62)	52 (88)	21 (64)	27	0.40
	Low-LC-high-CD	873 (38)	7 (12)	12 (36)	125	1.71
LCi_{final} & CDRAT	High-LC-low-CD	1440 (63)	51 (86)	22 (67)	28	0.43
	Low-LC-high-CD	847 (37)	8 (14)	11 (33)	106	1.38
	Full validation cohort	*2287 (100)*	*59 (100)*	*33 (100)*	*39*	*0.56*
NLST-eligible validation cohort (MILD)						
LCi_{final} & CD_{final}	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
LCi_{final} & CD_{CD}	High-LC-low-CD	1000 (82)	43 (91)	19 (76)	23	0.44
	Low-LC-high-CD	225 (18)	4 (9)	6 (24)	56	1.50
LCi_{final} & CD_{survey}	High-LC-low-CD	902 (74)	43 (91)	17 (68)	21	0.40
	Low-LC-high-CD	323 (26)	4 (9)	8 (32)	81	2.00
LCi_{final} & CDRAT	High-LC-low-CD	914 (75)	42 (89)	18 (72)	22	0.43
	Low-LC-high-CD	311 (25)	5 (11)	7 (28)	62	1.40
	NLST-eligible validation cohort	*1225 (100)*	*47 (100)*	*25 (100)*	*26*	*0.53*

CD, competing death; CD_{CT}, competing death CT model; CD_{final}, final competing death model; CD_{survey}, competing death survey model; CDRAT, Competing Death Risk Assessment Tool (10); High-LC-low-CD, risk group considered to have a relatively high LC risk and low risk of competing death (criteria described in the Figure 1 caption); LC, lung cancer; LCi_{final}, final lung cancer incidence model; Low-LC-high-CD, risk group considered to have a relatively low LC risk and high risk of competing death; MILD, Multicentric Italian Lung Detection; NLST, National Lung Screening Trial.

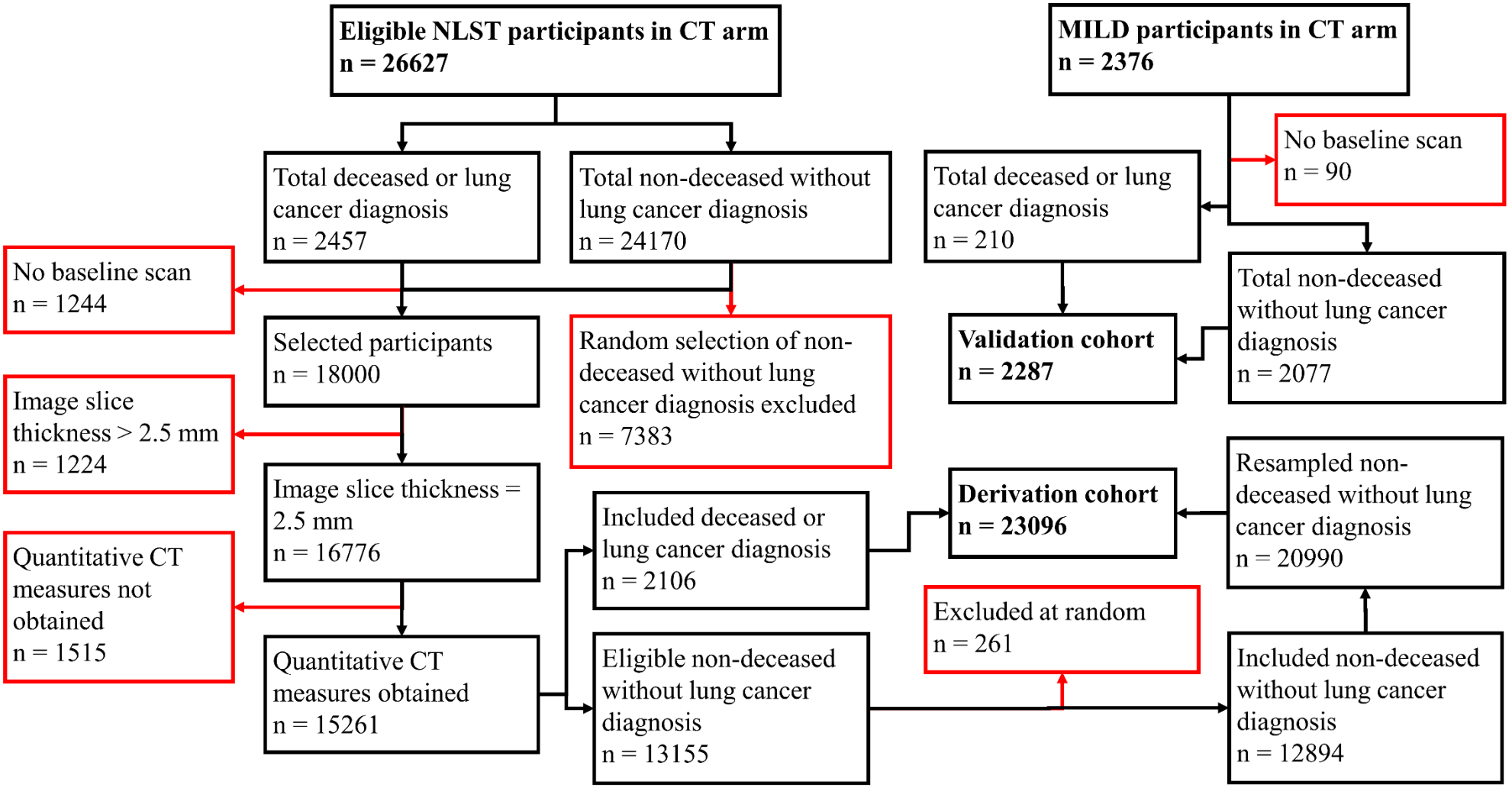

Figure S1: Cohort formation flowchart

For the derivation cohort, permission to use NLST was limited to up to 15000 participants. For this reason, all eligible participants who died or were diagnosed with lung cancer in the trial were included (n=2106) along with a random sample of eligible non-deceased participants without lung cancer (n=12894). To simulate the original CT cohort, the latter group was resampled without replacement to a total of 20990 non-deceased participants without lung cancer. In total, the derivation cohort consisted of 23096 participants. For the validation cohort (MILD), all eligible participants were included. MILD, Multicentric Italian Lung Detection; NLST, National Lung Screening Trial.

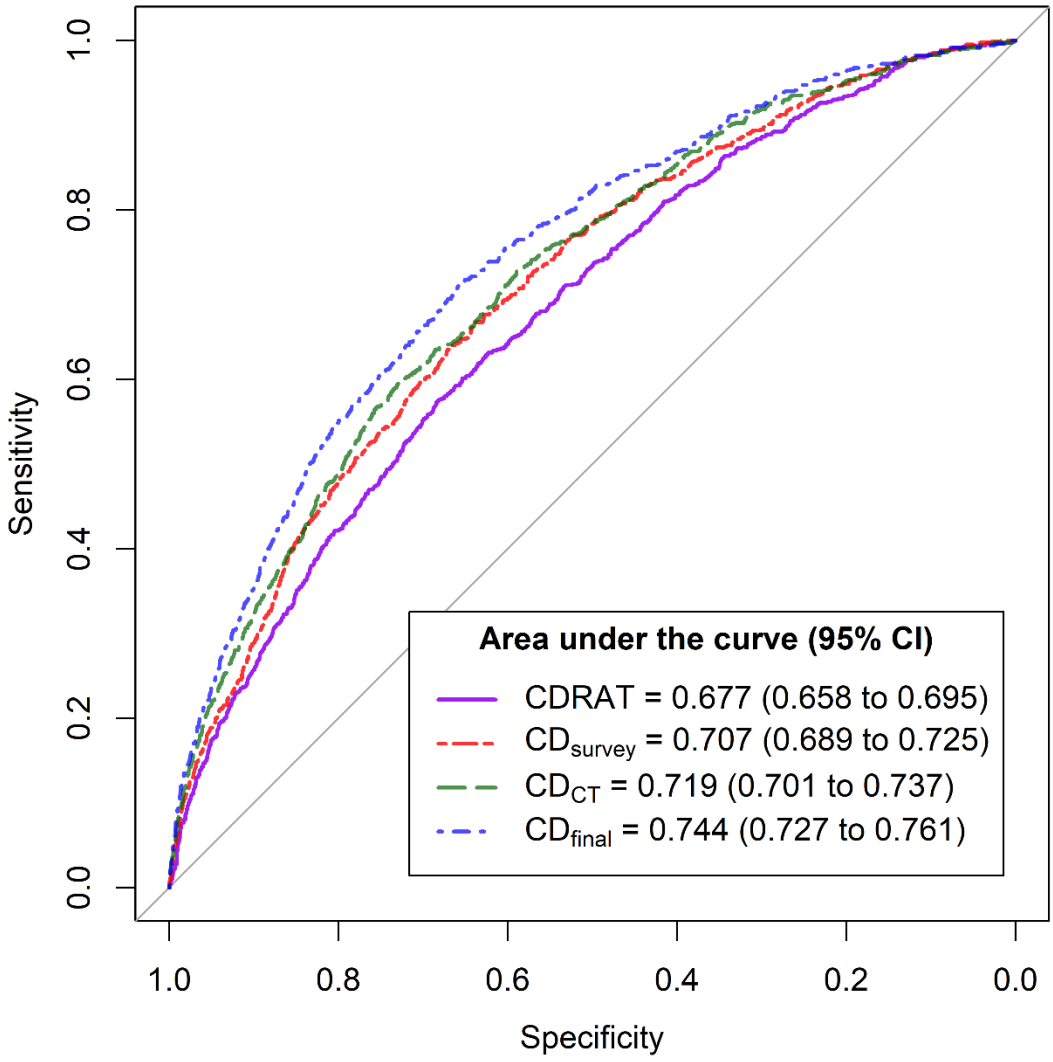

Figure S2: Receiver operating characteristic curves in the derivation cohort

CD_{CT} = CT-based competing death model; CD_{final}, final competing death model; CD_{survey}, self-reported patient characteristics-based competing death model; CDRAT, Competing Death Risk Assessment Tool; CI, confidence interval.

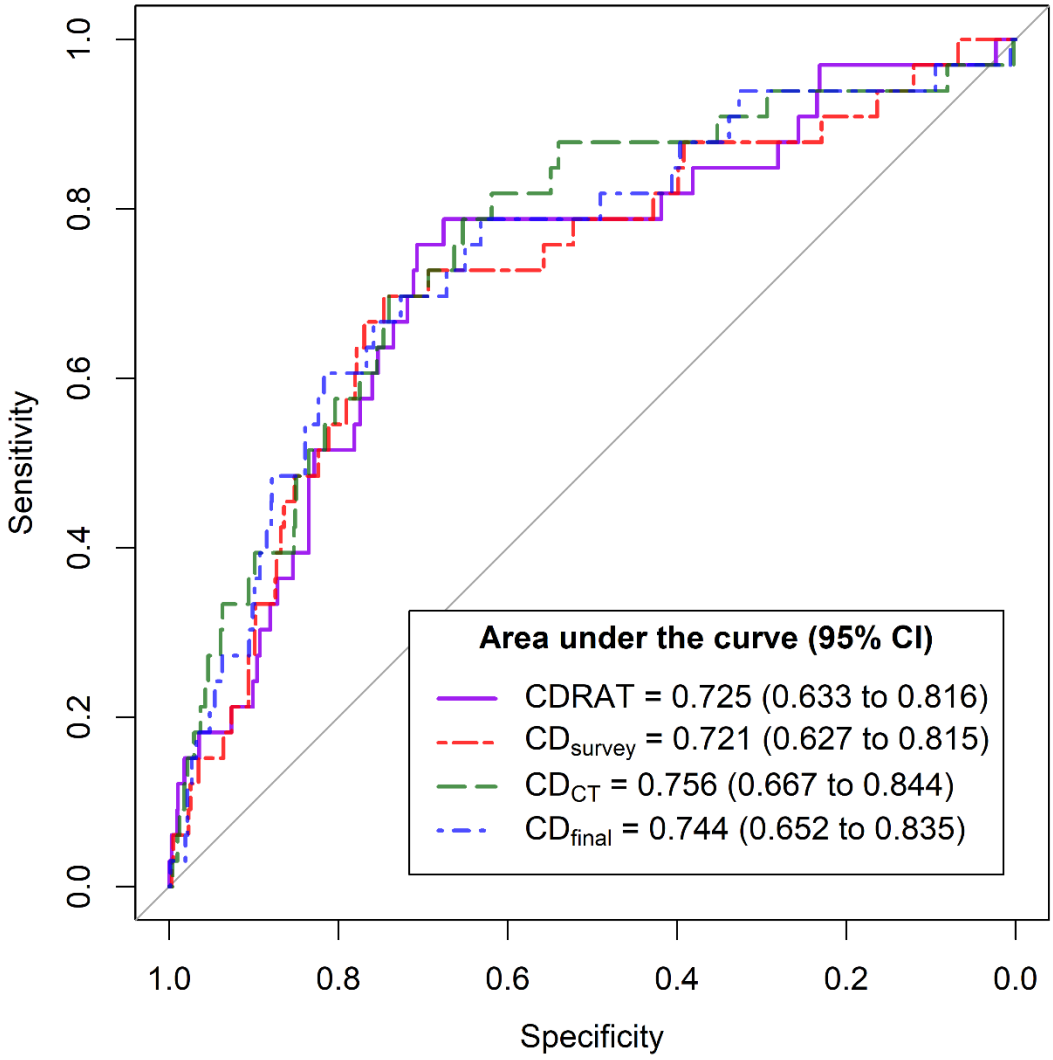

Figure S3: Receiver operating characteristic curves in the validation cohort

CD_{CT} = CT-based competing death model; CD_{final}, final competing death model; CD_{survey}, self-reported patient characteristics-based competing death model; CDRAT, Competing Death Risk Assessment Tool; CI, confidence interval.

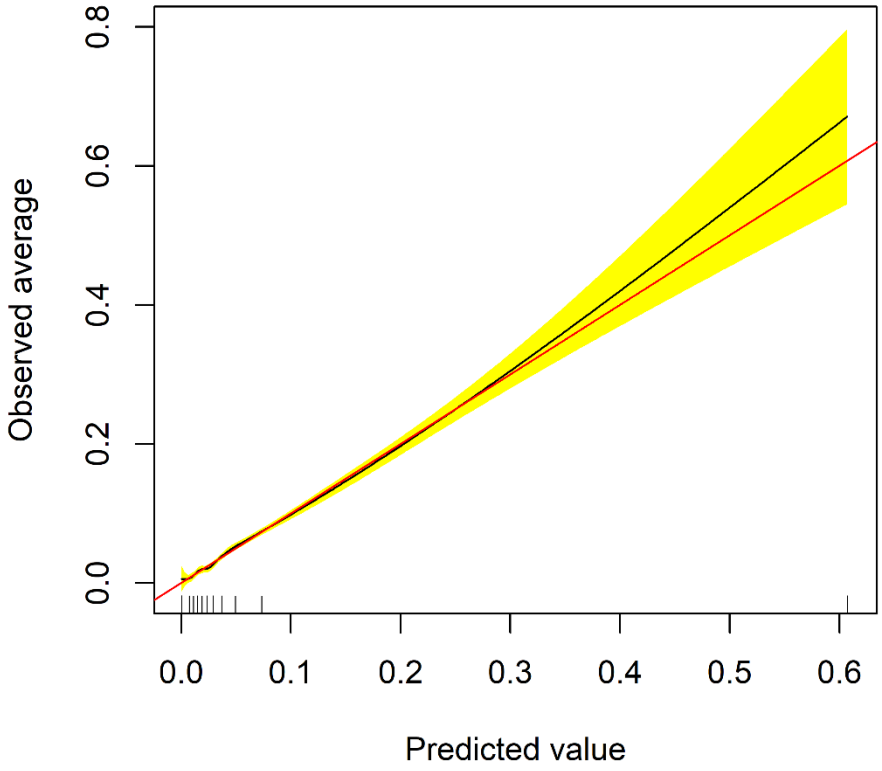

Figure S4: Calibration plot of CD_{final} in the derivation cohort

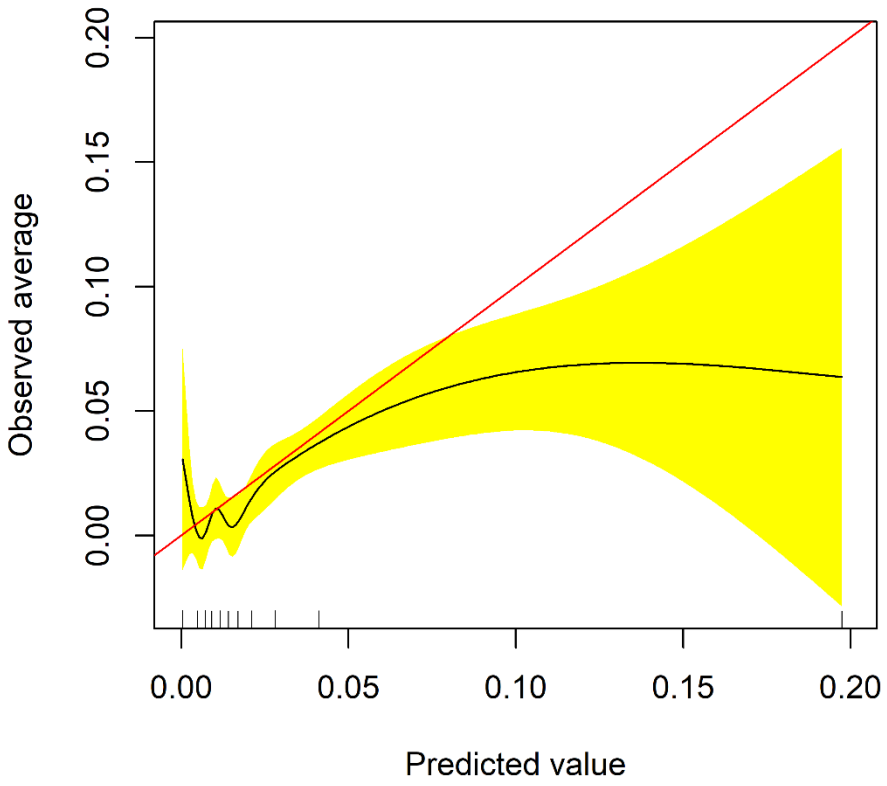

Figure S5: Calibration plot of CD_{final} in the validation cohort

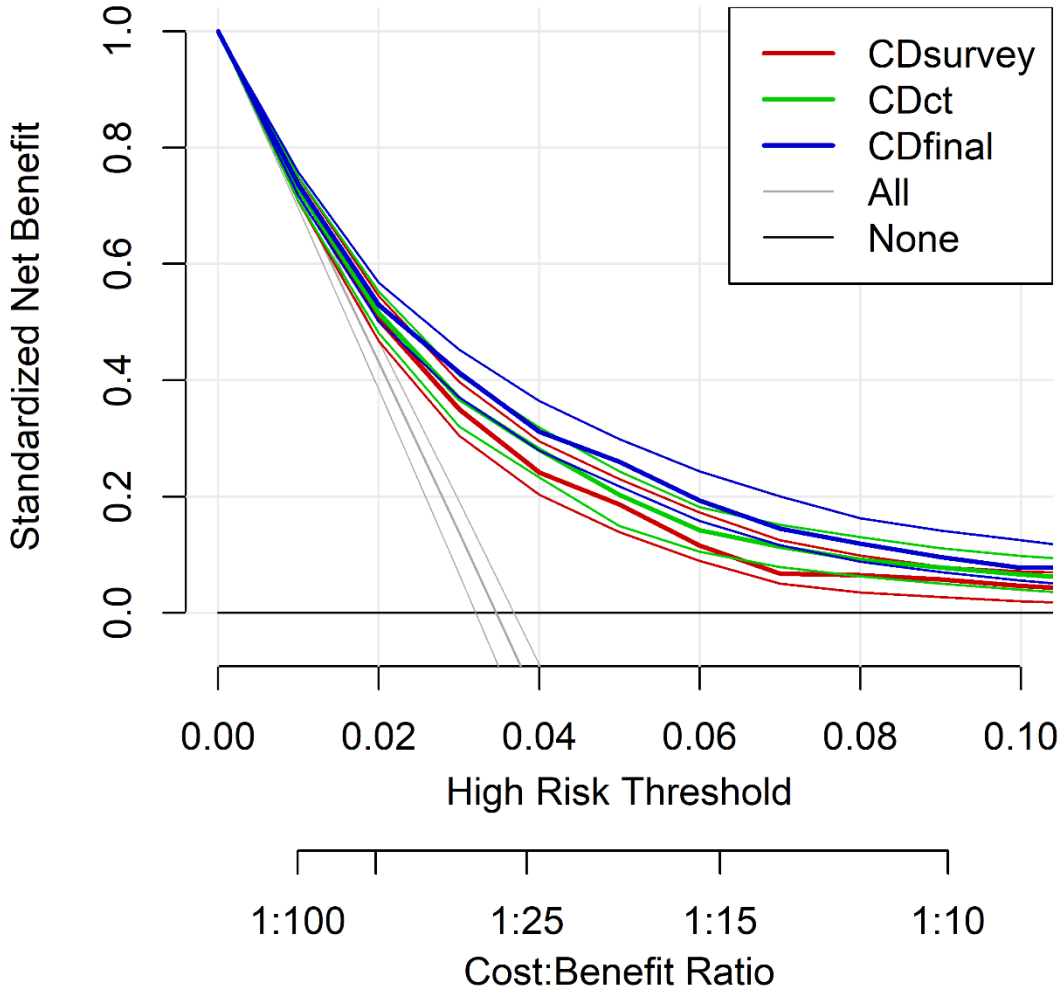

Figure S6: Decision curve analysis of competing death models in the derivation cohort

The thin lines accompanying each thicker line of the same color represent 95% confidence intervals (500 bootstrap samples). The colored lines represent the 5-year competing death models. The gray line ("All") represents the scenario where all participants are predicted to encounter a competing death. The horizontal black line at y-intercept zero represents the scenario where none are predicted to encounter the event. The x-axis represents the subjective preference threshold, where a higher risk threshold indicates a greater weight of false positive test findings per true positive test. The standardized net benefit is calculated as the difference between the true positive rate and the false positive rate (adjusted by the preference threshold and the event prevalence [5-year competing death]). A standardized net benefit greater than zero and greater than that of the gray "All" curve indicates that it would be beneficial to implement a model in practice. For a more detailed guide to the correct use and interpretation of decision curve analysis, please refer to Kerr et al. (11) or Vickers et al. (12).

CDct = CT-based competing death model; CDfinal, final competing death model; CDsurvey, self-reported patient characteristics-based competing death model.

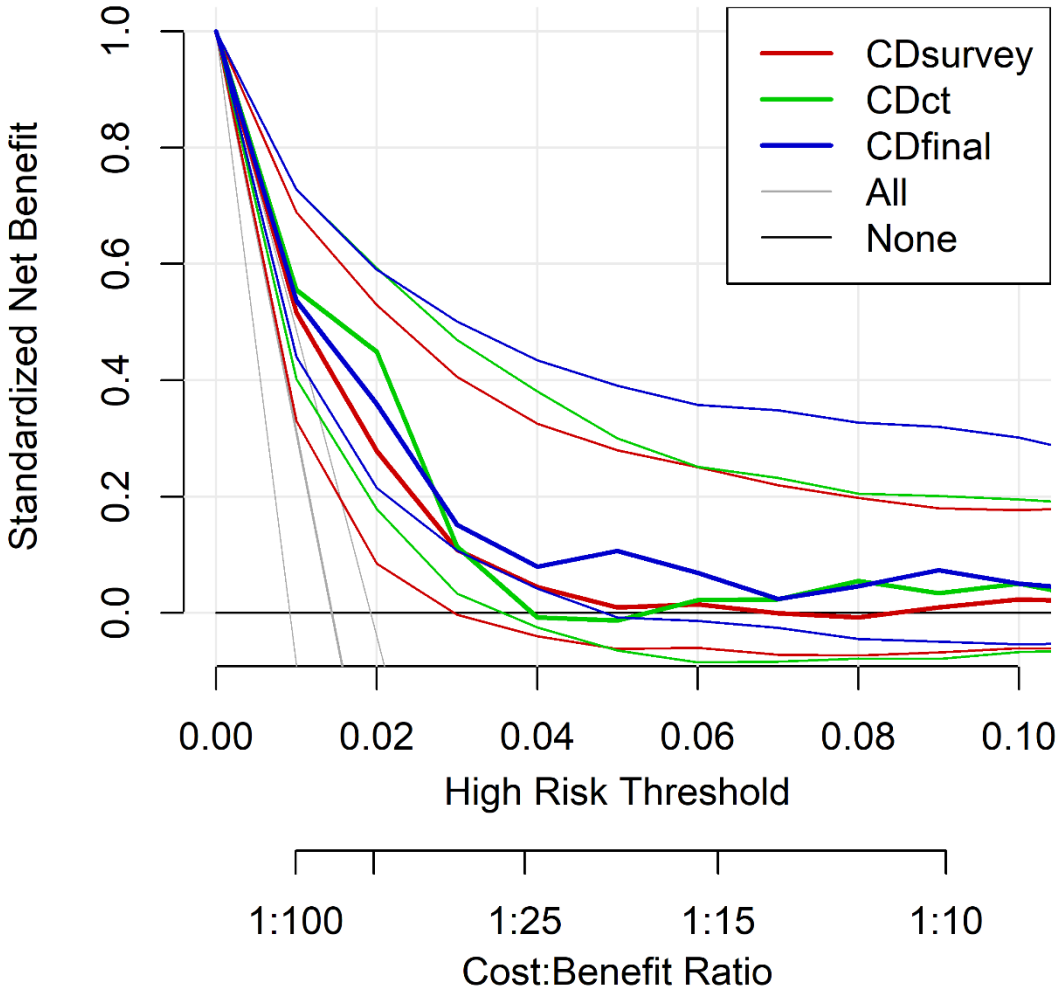

Figure S7: Decision curve analysis of competing death models in the validation cohort

The thin lines accompanying each thicker line of the same color represent 95% confidence intervals (500 bootstrap samples). The colored lines represent the 5-year competing death models. The gray line ("All") represents the scenario where all participants are predicted to encounter a competing death. The horizontal black line at y-intercept zero represents the scenario where none are predicted to encounter the event. The x-axis represents the subjective preference threshold, where a higher risk threshold indicates a greater weight of false positive test findings per true positive test. The standardized net benefit is calculated as the difference between the true positive rate and the false positive rate (adjusted by the preference threshold and the event prevalence [5-year competing death]). A standardized net benefit greater than zero and greater than that of the gray "All" curve indicates that it would be beneficial to implement a model in practice. For a more detailed guide to the correct use and interpretation of decision curve analysis, please refer to Kerr et al. (11) or Vickers et al. (12).

CDct = CT-based competing death model; CDfinal, final competing death model; CDsurvey, self-reported patient characteristics-based competing death model.

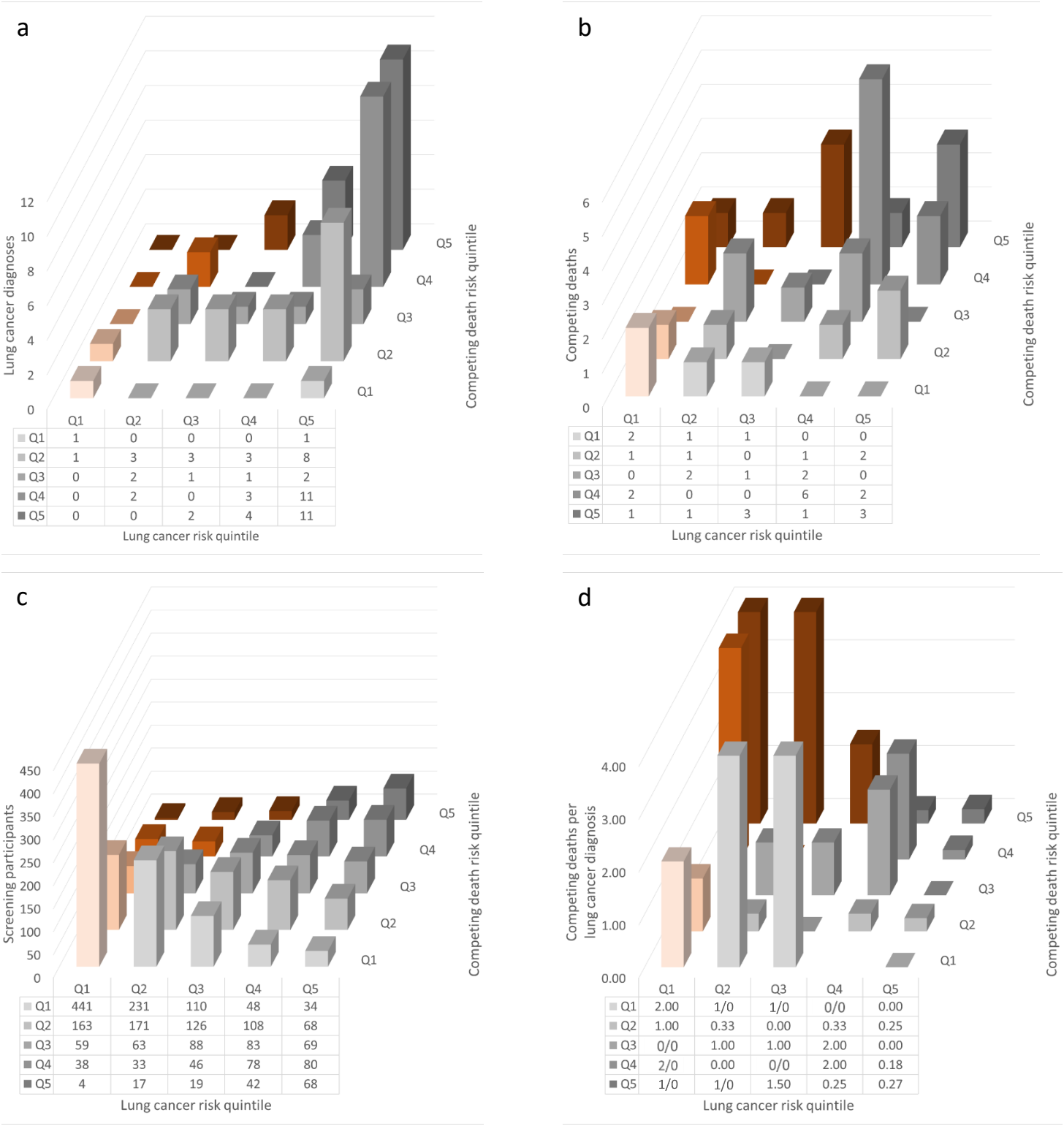

Figure S8: Outcomes per lung cancer and competing death risk quintile (validation cohort)

Three-dimensional column charts of (a) lung cancer diagnoses, (b) competing deaths, (c) screening participants, and (d) competing deaths per lung cancer diagnosis (y-axis truncated at 4). The lung cancer risk quintiles are 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 lung cancer risk and high risk of competing death (orange).

Q#, risk quintile where Q1 represents the lowest quintile and Q5 the highest.

References

1. National Lung Screening Trial Research Team, Aberle DR, Adams AM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *New England Journal of Medicine*. 2011;365(5):395–409. doi:10.1056/NEJMoa1102873.

2. Tammemägi MC, Katki HA, Hocking WG, et al. Selection criteria for lung-cancer screening. *The New England journal of medicine*. 2013;368(8):728–36. doi:10.1056/NEJMoa1211776.

3. Silva M, Schaefer-Prokop CM, Jacobs C, et al. Detection of Subsolid Nodules in Lung Cancer Screening. *Investigative Radiology*. 2018;53(8):441–9. doi:10.1097/RLI.0000000000000464.

4. Silva M, Prokop M, Jacobs C, et al. Long-Term Active Surveillance of Screening Detected Subsolid Nodules is a Safe Strategy to Reduce Overtreatment. *Journal of Thoracic Oncology*. 2018;13(10):1454–63. doi:10.1016/j.jtho.2018.06.013.

5. 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 Transactions on Medical Imaging*. 2018;37(2):615–25. doi:10.1109/TMI.2017.2769839.

6. Hoffmann U, Brady TJ, Muller J. Use of New Imaging Techniques to Screen for Coronary Artery Disease. *Circulation*. 2003;108(8). doi:10.1161/01.CIR.0000085363.88377.F2.

7. Gallardo-Estrella L, Lynch DA, Prokop M, et al. Normalizing computed tomography data reconstructed with different filter kernels: effect on emphysema quantification. *European Radiology*. 2016;26(2):478–86. doi:10.1007/s00330-015-3824-y.

8. Charbonnier J-PP, Pompe E, Moore C, et al. Airway wall thickening on CT: Relation to smoking status and severity of COPD. *Respiratory Medicine*. 2019;146:36–41. doi:10.1016/j.rmed.2018.11.014.

9. Schreuder A, Jacobs C, Lessmann N, et al. Combining pulmonary and cardiac computed tomography biomarkers for disease-specific risk modelling in lung cancer screening. *European Respiratory Journal*. 2021;2003386. doi:10.1183/13993003.03386-2020.

10. Katki HA, Kovalchik SA, Berg CD, Cheung LC, Chaturvedi AK. Development and Validation of Risk Models to Select Ever-Smokers for CT Lung Cancer Screening. *JAMA*. 2016;315(21):2300. doi:10.1001/jama.2016.6255.

11. Kerr KF, Brown MD, Zhu K, Janes H. Assessing the Clinical Impact of Risk Prediction Models With Decision Curves: Guidance for Correct Interpretation and Appropriate Use. *Journal of Clinical Oncology*. 2016;34(21):2534–40. doi:10.1200/JCO.2015.65.5654.

12. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. *Diagnostic and Prognostic Research*. 2019;3(1):18. doi:10.1186/s41512-019-0064-7.