Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: 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;369:245-54. DOI: 10.1056/NEJMoa1301851

Supplementary Appendix To

"Targeting Low-Dose CT Screening According to the Risk of Lung-Cancer Death"

Table Of Contents Page

Table S1. Frequency of coexisting pulmonary conditions at baseline in the National Lung Screening Trial (N=53,158)

Pulmonary Condition	Low-Dose CT (n=26,604)	Chest Radiography (n=26,554)
	Number (Percentage)	
Asbestosis	276 (1.0)	257 (1.0)
Bronchiectasis	853 (3.2)	899 (3.4)
Chronic Bronchitis	2,588 (9.7)	2,540 (9.6)
COPD	1,346 (5.1)	1,341 (5.1)
Emphysema	2,054 (7.7)	2,031 (7.6)
Fibrosis	70 (0.3)	58 (0.2)
Pneumonia	5,923 (22.3)	5,867 (22.1)
Sarcoidosis	48 (0.2)	49 (0.2)
Silicosis	30 (0.1)	27 (0.1)
Tuberculosis	278 (1.0)	293 (1.1)

COPD = Chronic obstructive pulmonary disease

Table S2. Internal and external validation of lung-cancer death risk prediction model for five-year lung-cancer death outcomes by NLST risk quintile group.

Five-Year Lung-cancer Death Risk (%)	Expected	Observed	Expected/ Observed	95% CI	AUC[a]	95% CI
Internal Validation[b]						
NLST Chest Radiography Group (n=26,554)						
>2.00	177.3	180	0.99	(0.85, 1.14)		
1.24-2.00	83.3	80	1.04	(0.84, 1.30)		
0.85-1.23	54.1	51	1.06	(0.81, 1.39)		
0.56-0.84	36.9	39	0.95	(0.70, 1.29)		
0.15-0.55	20.6	14	1.47	(0.87, 2.48)		
Overall	372.1	364	1.02	(0.92, 1.13)		
External Validation					0.72	(0.69, 0.75)
PLCO Chest Radiography Group, NLST Eligible (n=15,114)						
>2.00	122.0	137	0.90	(0.76, 1.05)		
1.24-2.00	51.7	59	0.88	(0.68, 1.13)		
0.85-1.23	32.3	40	0.81	(0.59, 1.10)		
0.56-0.84	19.4	14	1.40	(0.83, 2.37)		
0.15-0.55	8.9	6	1.49	(0.67, 3.32)		
Overall	234.6	256	0.92	(0.81, 1.04)	0.72	(0.69, 0.75)
PLCO Chest Radiography Group, NLST Ineligible Smokers Aged 55-74 Years[c] *(n=22,649)*						
>2.00	48.2	33	1.46	(1.04, 2.05)		
1.24-2.00	19.7	21	0.94	(0.61, 1.44)		
0.85-1.23	13.4	17	0.79	(0.49, 1.26)		
0.56-0.84	11.5	12	0.96	(0.54, 1.69)		
<0.56	5.1	5	1.02	(0.42, 2.45)		
Overall	97.9	88	1.11	(0.90, 1.37)	0.77	(0.72, 0.82)
PLCO Chest Radiography Group, NLST Eligible And Ineligible Smokers[c] *(n=37,763)*						
>2.00	127.1	142	0.90	(0.76, 1.06)		
1.24-2.00	63.2	71	0.89	(0.71, 1.12)		
0.85-1.23	45.7	57	0.80	(0.62, 1.04)		
0.56-0.84	39.5	35	1.13	(0.81, 1.57)		
<0.56	57.1	39	1.47	(1.07, 2.00)		
Overall	332.5	344	0.97	(0.87, 1.07)	0.80	(0.77, 0.82)

PLCO = Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial

a Area under the receiver operating characteristic curve, also known as the C-statistic

b Tenfold cross-validation: Overall Expected/Observed = 1.03 (95% CI, 0.90 to 1.12)

c 55-74 year-old current or former smokers

Table S3. Sample characteristics of external validation cohort, Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial randomized to the chest radiography group (n=37,763).

Characteristic	Smokers, Aged 55-74 Years	
	NLST Eligible[a]	NLST Ineligible[b]
	Number, Percentage	
N	15,114	37,763
Age, y		
55-59	5,631 (37.3)	13,587 (36.0)
60-64	4,890 (32.4)	11,904 (31.5)
65-69	3,159 (20.9)	8,146 (21.6)
70-74	1,434 (9.5)	4,126 (10.9)
Gender		
Male	9,216 (61.0)	21,953 (58.1)
Female	5,898 (39.0)	15,810 (41.9)
Body Mass Index		
<18.5	159 (1.1)	97 (0.3)
18.5-24.9	4,525 (29.9)	5,267 (13.9)
25-29.9	6,609 (43.7)	15,379 (40.7)
30-34.9	2,754 (18.2)	11,169 (29.6)
≥35	1,066 (7.1)	5,851 (15.5)
Family History		
None	13,289 (87.9)	33,483 (88.7)
One	1,644 (10.9)	3,894 (10.3)
Two or more	181 (1.2)	386 (1.0)
Emphysema		
No	13,868 (91.8)	36,086 (95.6)
Yes	1,246 (8.2)	1,677 (4.4)
Smoking status		
Former	9,008 (59.6)	29,823 (79)
Current	6,106 (40.4)	7,940 (21)
Pack-years		
<30	0 (0.0)	18,203 (48.2)
30-39.9	3,119 (20.6)	5,072 (13.4)
40-49.9	3,890 (25.7)	4,867 (12.9)
50-59.9	2,205 (14.6)	2,736 (7.2)
≥60	5,900 (39.0)	6,885 (18.2)
Quit years		
<1	6,631 (43.9)	8,959 (42.2)
1-4.9	2,105 (13.9)	3,274 (15.4)
≥5	6,378 (42.2)	9,008 (42.4)

a Current or former smokers (quit within 15 years), minimum 30 pack-years
b Current or former smokers

Table S4. Number of prevented lung-cancer deaths by pulmonary comorbidity status (per 10,000 person-years) and five-year risk of lung-cancer death.

Lung-cancer Death Risk (%)	Preexisting Pulmonary Conditions[a]					
	None		One		Two or More	
	No. (%)	Risk Difference	No. (%)	Risk Difference	No. (%)	Risk Difference
Overall	34,280 (100)	6.2 (1.7, 10.7)	13,150 (100)	9.6 (1.5, 17.7)	5,728 (100)	-0.5 (-15.4, 14.3)
>2.00	5482 (16.0)	8.5 (-10.2, 27.2)	3104 (23.6)	20.2 (-6.0, 46.3)	2046 (35.7)	8.6 (43.0, -25.8)
1.24-2.00	6601 (19.3)	14.7 (3.2, 26.2)	2744 (20.9)	6.5 (-11.0, 24.1)	1286 (22.5)	1.0 (-26.5, 28.4)
0.85-1.23	7160 (20.9)	5.8 (-2.6, 14.1)	2527 (19.2)	9.6 (-4.4, 23.5)	945 (16.5)	-11.6 (-39.3, 14.0)
0.56-0.84	7367 (21.4)	3.2 (-3.8, 10.2)	2448 (18.6)	9.9 (-3.2, 23.0)	816 (14.2)	-12.6 (-39.3, 14.2)
0.15-0.55	7670 (22.4)	1.8 (-3.2, 6.7)	2327 (17.6)	-4.7 (4.5, -13.9)	635 (11.1)	0.2 (-22.2, 22.8)
Linear Trend						
Est. (95% CI)		2.9 (-0.1, 5.9)		5.2 (0.3, 10.0)		1.8 (-6.9, 10.5)
P-value		0.06		0.03		0.68

a Total lifetime diagnoses of asbestosis, bronchiectasis, chronic bronchitis, COPD, emphysema, lung fibrosis, lung sarcoidosis, pneumonia, silicosis, and tuberculosis.

Table S5. Screening outcomes by round and five-year risk of lung-cancer death in the Low-Dose CT screening group.

	Five-Year Risk of Lung-cancer Death (%)					
Round	0.15-0.55	0.56-0.84	0.85-1.23	1.24-2.00	>2.00	Total
T0						
True Positive	28 (0.5)	41 (0.8)	66 (1.2)	80 (1.5)	162 (3.1)	377 (1.4)
False Positive	1,175 (22.3)	1,281 (24.1)	1,343 (24.9)	1,443 (27.2)	1,563 (29.4)	6,805 (25.6)
Negative	4,004 (75.9)	3,941 (74.2)	3,921 (72.7)	3,717 (69.9)	3,499 (65.9)	19,802 (71.7)
Not Screened[a]	69 (1.3)	47 (0.9)	66 (1.2)	74 (1.4)	84 (1.6)	340 (1.3)
T1						
True Positive	21 (0.4)	32 (0.6)	46 (0.9)	83 (1.6)	122 (2.3)	304 (1.1)
False Positive	1,197 (22.7)	1,215 (2.9)	1,334 (24.7)	1,363 (25.6)	1,491 (28.1)	6,600 (24.8)
Negative	3,798 (72.0)	3,759 (70.8)	3,634 (67.3)	3,454 (65.0)	3,164 (59.6)	17,809 (67.0)
Not Screened	260 (4.9)	304 (5.7)	382 (7.1)	414 (7.8)	531 (10.0)	1,891 (7.1)
T2						
True Positive	20 (0.4)	23 (0.4)	37 (0.7)	55 (1.0)	86 (1.6)	221 (0.8)
False Positive	630 (11.9)	707 (13.3)	782 (14.5)	795 (15.0)	911 (17.2)	3,825 (14.4)
Negative	4,302 (81.5)	4,181 (78.7)	4,078 (75.6)	3,920 (73.8)	3,571 (67.3)	20,052 (75.4)
Not Screened	324 (6.2)	399 (7.6)	499 (9.2)	544 (10.2)	740 (13.9)	2,506 (9.4)
Overall[b]						
True Positive	51 (1.0)	73 (1.4)	113 (2.2)	150 (2.9)	280 (5.3)	667 (2.5)
False Positive[c]	1,648 (31.2)	1,806 (34.0)	1,911 (35.4)	1,973 (37.1)	2,146 (40.4)	9,484 (35.6)
Negative	3,534 (66.9)	3,405 (64.1)	3,336 (61.8)	3,147 (59.2)	2,833 (53.4)	16,255 (61.1)
Not Screened	43 (0.9)	26 (0.5)	36 (0.7)	44 (0.8)	49 (0.9)	198 (0.8)

T0 = First Screening Round, T1 = Second Screening Round, T2 = Third Screening Round

a Not screened includes missed screens, unreadable screens, and screens not performed due to lung cancer diagnosis.

b Each person was given a screening status based on the outcomes of the three screening rounds. Persons with any true positive result were classified as true positives; persons with any positive result but no true positive result were classified as false positives; persons with any screen result but no positive result were classified as negatives; persons with no screening result were considered not screened.

c At least one positive screen without any linked lung cancer diagnosis.

Linked-Year Method for Classification of Screen-Detected Lung Cancers

Screen Link Description

The classification of screen-detected lung cancers was based on the results of the diagnostic follow-up occurring within one year of a linked screen. A **screen link** started at the screen (T0, T1, or T2) and extended forwards from the end of the current diagnostic chain to the next event. An event could be another procedure, a lung cancer diagnosis, or another screen. If there was no next event, the next event was a screen, or the next event occurred more than 12 months after the current end of the chain, the chain ended. If the next event was a lung cancer within one year, then that cancer was considered screen-linked. Otherwise the next event was another procedure within one year, and that became the new end of the chain and the process repeated.

Because the next event could start at a maximum of 12 months from a screen and a screen-linked diagnosis at most 12 months from the "next event", the linked-year method could include a lung cancer diagnosis that occurred as much as 24 months from the initiating screen.

Event Data Collection

Procedure and lung cancer information was captured on the Diagnostic Evaluation form. A Diagnostic Evaluation form was needed when there was a positive screen or a report of a lung cancer diagnosis. If a participant had a positive screen, the study attempted to collect medical records for any follow-up the participant might have sought. Information on follow-up procedures occurring up to one year after the screen was collected, or up to two years for nodules found on T2 screens that were either newly detected or showed growth from previous screens. Follow-up beyond one year (or two years for T2 screens) could be collected if the screening center determined that the follow-up was prompted by the screen. If the trial learned of a lung cancer diagnosis not resulting from an NLST screen (usually from either a participant self-report on the Annual Study Update or from a death certificate), the trial attempted to collect records back to whatever non-trial exam or initial presentation with symptoms led to the diagnosis.

Risk Prediction Models

Five measures of risk were considered in our assessment of risk-based variation in Low-Dose CT screening efficacy. Below we describe the interpretation and computation of each measure.

1. Lung-cancer Death Risk Measure

Interpretation: Five-year cumulative risk of dying of lung cancer from the time of randomization in the absence of Low-Dose CT screening.

Computation: We developed the lung-cancer death model with data from the chest radiography group of the NLST. The formula for the cumulative lung-cancer death risk is

$$R_1 = \int_0^5 \lambda_1(u; x_1) \exp\{-\int_0^u (\lambda_1(v; x_1) + \lambda_2(v; x_2))dv\},$$

where $\lambda_1(t)$ and $\lambda_2(t)$ are the cause-specific hazard models for lung-cancer death and all other causes of death, respectively. Breslow's semiparametric estimator was used to estimate the baseline hazard for each event type given the observed death in the chest radiography group of the NLST.

Each cause-specific hazard followed Cox's proportional hazards model. For the hazard model for lung-cancer death, we used multivariable regression procedures to identify significant risk factors among a set of candidate predictive factors, which included age, gender, race, body mass index, family history of lung cancer, previous or current occupation with dust exposure, previous or current occupation with asbestos exposure, diagnosis of emphysema, diagnosis of COPD, number of lung comorbidities (categorized as none, one, two or more as in Table 1 of the manuscript), pack-years smoked, average cigarettes smoked per day, and years since quitting. For continuous variables, we fit stratified Kaplan-Meier lung cancer survival curves for each quintile of the continuous variable to guide the representation for the continuous variables, i.e. whether to treat the variable as categorical and which categories to assign. We fit a model with all main effects and retained any effect that was significant at the 5% level and that was also independently selected in a lasso regression with penalty parameter determined by tenfold cross-validation.[1] We checked for interactions with gender among the selected risk factors and found no interactions at the 5% level of significance. We also assessed the proportional hazards assumption for each factor by examining the correlation of the survival times and the scaled Schoenfeld residuals.

For the competing death hazard model, we fit a model with all the predictive factors of the lung-cancer death hazard model and, if not selected risk factors, age, gender, and race. As with the lung-cancer death hazard model, we retained all factors that were significant at the 5% significance level, and we evaluated proportionality in the selected model based on Schoenfeld residuals.

The selected risk factors x_1 and x_2 and their corresponding hazard ratios are listed in Table 1. There was no evidence of non-proportionality for any factor in the lung-cancer death hazard model or the competing death model.

The developed absolute risk model can be used to estimate the absolute risk of lung-cancer death for any specified time from randomization. We chose the five-year projection interval since the median time of follow-up for the January 15, 2009 cutoff date was 5.5 years.

2. Bach 2003 Lung Cancer Risk Measure

Interpretation: One-year cumulative risk of lung cancer.

Computation: Bach and colleagues developed a risk prediction model based on data from the CARET trial.[2] The model risk factors included gender, age, asbestos exposure, smoking duration, quitting duration, and average cigarettes smoker per day. Cubic splines with knots at the quartiles of the sample values were used for continuous variables. The model coefficients are tabulated in Table S6. Given the relative risks β, risk factors *x*, and the baseline probability of being free of lung cancer after one-year equal to $S_0 = 0.99629$, the predicted lung cancer risk is

$$R_2 = 1 - {S_0}^{\exp(\beta' x)}.$$

Table S6. Risk factors and log relative risk estimates for the Bach 2003 lung-cancer incidence model.

Risk Factor	Coding	Log Relative Risk (β)
Intercept		-9.7960571
Female	Binary	-0.0582726
Asbestos Exposure	Binary	0.2153936
Age	Age (Continuous)	0.0703228
	$I(Age>53)(Age - 53.5)^3$	-0.0000938
	$I(Age>61)(Age - 61.9)^3$	0.0001828
	$I(Age>70)(Age - 70.9)^3$	-0.0000890
Quit Duration	Years (Continuous)	-0.0856848
	$Years^3$	0.0065500
	$I(Years>0)(Years - 0.5051)^3$	-0.0068306
	$I(Years>12)(Years - 12.296)^3$	0.0002806
Smoke Duration	Years (Continuous)	0.1142530
	$I(Years>27)(Years - 27.66)^3$	-0.0000801
	$I(Years>40)(Years - 40)^3$	0.0001707
	$I(Years>50)(Years - 50.9)^3$	-0.0000906
Cigarettes Per Day	Rate (Continuous)	0.0608184
	$I(Rate>15)(Rate - 15)^3$	-0.0001465
	$I(Rate>20)(Rate - 20.19)^3$	0.0001849
	$I(Rate>40)(Rate - 40)^3$	-0.0000383

I(x) takes the value one if condition *x* is true and zero otherwise.

3. Spitz 2007 Lung Cancer Risk Measure

Interpretation: One-year cumulative risk of lung cancer.

Computation: The lung cancer model of Spitz and colleagues was developed from a case-control study of individuals living in Texas. Separate models were generated for never, former, and current smokers.[3] The risk factors and estimated log odds ratios for the former- and current-smoker models are given in Table S7. The one-year cumulative risk of lung cancer was computed as

$$R_3 = \frac{h_{1j(i)k(i)s(i)} r_{ik(i)}}{(h_{1j(i)k(i)s(i)} r_{ik(i)} + h_{2j(i)s(i)})}[1\text{-exp}\{-(h_{1j(i)k(i)s(i)} r_{ik(i)} + h_{2j(i)s(i)})\}],$$

where *i* indexes an individual, *j(i)* is an index for age, *k(i)* for smoking status and *s(i)* gender. The quantity r_{ik} is the individual's relative risk,

$$r_{ik} = \exp(\beta_k{}' x_i),$$

given the log odds ratios βk for the *kth* smoking-status model. The quantity h_{1jks} is the one-year age, smoking-status, and sex-specific baseline hazard rate for lung cancer, while h_{2js} is age- and sex-specific one-year death rate.

To obtain the age-, smoking- status-, and gender-specific baseline hazard rate for lung cancer incidence, the authors used the relationship between the attributable risk and baseline incidence, correcting the incidence for the prevalence of never, former, and current smokers among lung cancer patients. Let AR_k be the attributable risk for the *kth* smoking model. The attributable risk is computed by taking the inverse sum of the relative risk among cases of the model. Suppose the number of cases was *mk*, then

$$AR_k = \left(1 - \frac{1}{m_k \sum_{I=1}^{m_k} \frac{1}{r_{ik}}}\right).$$

In the Spitz case-control data set, the attributable risk for current smokers was 0.51404 and for former smokers was 0.45352. Given the age- and sex-specific lung cancer incidence I_{js}, the smoking-status-specific incidence was computed as $v_{jks} = c_{ks} I_{js}$, where c_{ks} is a constant factor that depends on smoking and gender. These quantities yield the baseline lung cancer incidence estimate,

$$h_{jks} = v_{jks}(1 - AR_k).$$

The authors used the SEER incidence rates in Table S8 for I_{js}, and the adjustment factors c_{ks} are 3.17 for male former smokers, 3.88 for male current smokers, 3.76 for female former smokers and 4.17 for female current smokers.

Table S7. Risk factors and log relative risk estimates for the Spitz 2007 lung cancer incidence model.

	Log Relative Risk	
Risk Factors	Former Smoker	Current Smoker
Emphysema	0.9734	0.9734
Dust Exposure	0.4654	0.3067
Asbestos Exposure		0.4109
Family History		
Any Cancer (2 or more)	0.4636	
Smoking-Related Cancer (1 or more)		0.3859
No Hay Fever	0.3711	0.4047
Age Quit		
<42	1.0000	
42-53	0.2130	
53+	0.4080	
Pack-years		
<28		1.0000
28-41		0.2219
42-57.4		0.3747
57.5 or more		0.6151

Table S8. SEER 2005 lung cancer incidence and NCHS 1999-2003 mortality rates per 100,000 person-years.

	Male		Female	
Ages	Incidence	Mortality	Incidence	Mortality
40-44	10.8	275.1	11.0	153.2
45-49	25.5	400.7	23.2	218.8
50-54	56.6	560.0	45.5	313.4
55-59	116.6	786.9	93.9	479.1
60-64	221.2	1,210.2	164.9	762.9
65-69	346.8	1,855.1	246.9	1,197.0
70-74	478.1	2,947.4	318.7	1,968.3
75-79	564.4	4,836.4	344.7	3,306.1
80-84	532.4	7,980.7	308.3	5,761.2
85+	498.4	15,559.4	266.7	14,016.2

4. LLP 2008 Lung Cancer Risk Measure

Interpretation: Five-year cumulative risk of lung cancer.

Computation: Cassidy and authors developed a prediction model for lung cancer risk with data from the Liverpool Lung Project, a case-control study of lung cancer cases among residents of Liverpool.[4] Given the risk log odds ratios β and risk factors *x*, the risk of lung cancer according to the LLP model is

$$R_4 = \frac{\exp(\beta' x)}{1 + \exp(\beta' x)}.$$

The risk factors and log odds ratios are listed in Table S9. The intercept term of the model was adjusted so that the mean value of risks among controls was equal to the age- and gender-specific lung-cancer incidence rates in the population. The adjustment values are provided in Table S10. To calculate the baseline incidence for an individual of a given age, the authors interpolated between values of adjacent 5-year age ranges.

Table S9. Risk factors and log relative risk estimates for the LLP 2008 lung cancer incidence model.

Risk Factors	Log Odds Ratio
Pneumonia	0.620
Asbestos Exposure	0.634
Prior Cancer	0.675
Family History	
None	1.000
Early Onset	0.703
Late Onset	0.168
Smoking Duration	
Never	1.000
1-20 years	0.769
21-40 years	1.452
41-60 years	2.507
60 years or more	2.724

Table S10. Lung cancer incidence rates per 100,000 person-years from Liverpool cancer registries 2002-2004 and corresponding adjusted intercept for logistic model.

	Male		Female	
Ages	Incidence	Intercept	Incidence	Intercept
40-44	15.5	-9.06	6.0	-9.90
45-49	37.9	-8.16	37.3	-8.06
50-54	88.7	-7.31	68.1	-7.46
55-59	172.3	-6.63	175.2	-6.50
60-64	329.0	-5.97	230.6	-6.22
65-69	487.4	-5.56	288.1	-5.99
70-74	616.5	-5.31	465.0	-5.49
75-79	950.6	-4.83	594.2	-5.23
80-84	1096.4	-4.68	497.1	-5.42

5. Tammemagi 2011 Lung Cancer Risk Measure

Interpretation: Nine-year cumulative risk of lung cancer.

Computation: Tammemagi and authors developed a lung cancer incidence model for smokers with data from the Prostate, Lung, Colorectal and Ovarian Trial.[5] Like the LLP measure, the risk score is based on a logistic regression model and is computed as

$$R_5 = \frac{\exp(\beta' x)}{1 + \exp(\beta' x)},$$

given the risk factors and log odds ratios listed Table S11. The continuous variables age, pack- years, and quit duration were modeled with restricted cubic splines. The knots for age were 55, 60, 64, and 62 years; for pack-years 3.25, 23.25, and 63; for quitting time 0, 15, and 35 years.

Table S11. Risk factors and log odds ratios for the Tammemagi 2011 lung cancer incidence model.

Risk Factor	Coding	Log Odds Ratios (β)
Intercept		-18.110
Education	Trend (7 levels)	-0.075
Body Mass Index	Continuous	-0.028
Family History of Lung Cancer	Binary	0.445
COPD	Binary	0.318
Chest Radiography, Past 3 yrs		0.111
Current Smoker		0.305
Age	Spline 1	0.219
	Spline 2	-0.350
	Spline 3	0.791
Quit Duration	Spline 1	-0.057
	Spline 2	0.046
Smoke Duration	Years	0.012
Pack-years	Spline 1	0.057
	Spline 2	-0.052

Validation

The external validation of the lung cancer death model was performed with the individuals in the chest radiography arm of the PLCO who were current or former smokers between the ages of 55 and 74 years at randomization. The expected lung cancer deaths within *t* years of randomization were computed from the prediction model developed from the NLST chest radiography group outcomes. Let $r(t)_i$ be the predicted risk for the *ith* smoker in the PLCO chest radiography group, $y(t)_i$ be the indicator of lung cancer death within *t* years, and δ_i be the indicator of NLST eligibility. The expected deaths for ever smokers were computed as $E_t = \sum_i r(t)_i$ and for NLST-eligible smokers $\mathcal{E}_t = \sum_i r(t)_i \delta_i$, with the corresponding observed counts $O_t = \sum_i r(t)_i$ and for NLST-eligible smokers $\mathrm{O}_t = \sum_i r(t)_i \delta_i$.

References

1. Tibshirani R. The lasso method for variable selection in the Cox model. Stat Med 1997;16:385-95.

2. Bach PB, Kattan MW, Thornquist MD, Kris MG, Tate RC, Barnett MJ, Hsieh LJ, Begg CB. Variations in lung cancer risk among smokers. J Natl Cancer Inst 2003; 95(6):470-8.

3. Spitz MR, Hong WK, Amos CI, Wu XF, Schabath MB, Dong Q, Shete S, Etzel CJ. A risk model for prediction of lung cancer. J Natl Cancer Inst 2007; 99(9):715-26.

4. Cassidy A, Myles JP, van Tongeren M, Page RD, Liloglou T, Duffy SW, Field JK. The llp risk model: an individual risk prediction model for lung cancer. Br J Cancer 2008; 98(2):270-6.

5. Tammemagi CM, Pinsky PF, Caporaso NE, et al. Lung cancer risk prediction: Prostate, lung, colorectal and ovarian cancer screening trial models and validation. J Natl Cancer Inst 2011;103(13):1058-68.