IJC
International Journal of Cancer

Small median tumor diameter at cure threshold (<20 mm) among aggressive non-small cell lung cancers in male smokers predicts both chest X-ray and CT screening outcomes in a novel simulation framework

Deborah L. Goldwasser and Marek Kimmel

Department of Statistics, Rice University, Houston, TX

The effectiveness of population-wide lung cancer screening strategies depends on the underlying natural course of lung cancer. We evaluate the expected stage distribution in the Mayo CT screening study under an existing simulation model of non-small cell lung cancer (NSCLC) progression calibrated to the Mayo lung project (MLP). Within a likelihood framework, we evaluate whether the probability of 5-year NSCLC survival conditional on tumor diameter at detection depends significantly on screening detection modality, namely chest X-ray and computed tomography. We describe a novel simulation framework in which tumor progression depends on cellular proliferation and mutation within a stem cell compartment of the tumor. We fit this model to randomized trial data from the MLP and produce estimates of the median radiologic size at the cure threshold. We examine the goodness of model fit with respect to radiologic tumor size and 5-year NSCLC survival among incident cancers in both the MLP and Mayo CT studies. An existing model of NSCLC progression under-predicts the number of advanced-stage incident NSCLCs among males in the Mayo CT study (p-value = 0.004). The probability of 5-year NSCLC survival conditional on tumor diameter depends significantly on detection modality (p-value = 0.0312). In our new model, selected solution sets having a median tumor diameter of 16.2–22.1 mm at cure threshold among aggressive NSCLCs predict both MLP and Mayo CT outcomes. We conclude that the median lung tumor diameter at cure threshold among aggressive NSCLCs in male smokers may be small (<20 mm).

Introduction

The first randomized clinical trial (RCT) to demonstrate a mortality benefit associated with screening for lung cancer is the national lung screening trial (NLST), reporting a 20.0% reduction in lung cancer deaths among individuals screened annually by computed tomography (CT) relative to individuals screened annually by chest X-ray.[1] The prostate, lung, colorectal, and ovarian (PLCO)[2] study reported no mortality benefit associated with screening by chest X-ray after 13.5 years of follow-up, consistent with earlier findings from the Mayo Lung Project (MLP).[3]

Key words: lung cancer, simulation, mathematical model, tumor size, doubling times

Grant sponsor: National Cancer Institute CISNET; **Grant number:** U01CA097431B; **Grant sponsor:** National Science Foundation VIGRE fellowship; **Grant number:** DMS-0739420

DOI: 10.1002/ijc.27599

History: Received 26 Nov 2011; Accepted 27 Mar 2012; Online 17 Apr 2012

Correspondence to: Deborah L. Goldwasser, Department of Statistics, Rice University, 6100 Main Street, MS-138, Houston, TX 77005, Tel.: 713-348-3264, Fax: 713-348-5476, E-mail: deborah.l.goldwasser@rice.edu

The NLST and PLCO results suggest that screening for lung cancer by CT is likely to be effective in reducing lung cancer mortality whereas screening by chest X-ray will not. The NLST study found that CT detects lung nodules larger than 4mm at a three-fold higher rate than chest-X-ray.[1] However, the PLCO and MLP results likewise demonstrated an increased rate of early-stage lung cancer detection attributable to chest X-ray screening, relative to usual care.[2,4] Debate over the interpretation of chest X-ray screening results persists.[5–8] Estimates of the median size at which lung cancers transition to advanced-stage of 40 mm suggest that a mortality benefit to chest X-ray screening should exist.[9]

The failure of chest X-ray screening studies to produce a mortality benefit has been widely attributed to overdiagnosis.[6] However, data suggests that screen-detected lung cancers typically demonstrate rapid growth. A systematic literature review of volume doubling times (VDTs) reports that the mean proportions of VDTs greater than 400 days among lung cancers detected by routine care, chest X-ray and CT are 3, 8, and 27%, respectively.[10] Average proportions of tumors with VDTs less than 100 days among lung cancers detected by routine care, chest X-ray and CT are 44, 39, and 29%, respectively.[10] The proportion of rapidly-growing lung cancers may in fact be higher. Inclusion of baseline cancers (influenced by length bias) and exclusion of single

measurement cancers may bias VDT estimates toward slower growth.[11] In the Mayo CT, detection of baseline cancers was two-fold higher in females than males, suggesting that the proportion of rapid-growing cancers may be higher in men than women.[11] Furthermore, nonsolid or part-solid nodules may appear slow-growing but in fact become denser and increase in cell volume.[12]

The effectiveness of lung cancer screening depends on the underlying natural course of lung cancer. In this study, we characterize a model of lung cancer progression having two key features, namely VDT and the radiologic size at cure threshold, the size at which a lung cancer is no longer curable. When lung cancer is detected at a localized stage, it can be successfully treated by surgical resection followed by chemotherapy.

We fit a model of lung cancer natural course to randomized trial data from the MLP. MLP data distinguishes between lung cancers detected during the study and lung cancers detected at the time of the prevalence screen. Prevalent cancers are subject to length bias since slow-growing or indolent cancers are likely to be over-represented in a population at a given point in time. In contrast, incident cancers arise during the screening study, following a negative prevalence screen. In the context of the MLP, incident cancers from the usual-care arm and the chest X-ray arm provide nearly unbiased random samples with respect to lung cancer growth and progression. On average, incident cancers in the usual-care arm are expected to reflect a later progression point in the natural course of lung cancer.

Evidence from the surveillance, epidemiology, and end results (SEER) cancer registry suggests that the smaller the tumor size, the more likely it is that a lung cancer has not spread to lymph nodes or metastasized. In a study of 84,152 non-small cell lung cancer (NSCLC) cases documented in SEER, among tumors less than 15 mm in diameter, the proportion of Stage I cancers was 54% (N =7,327) whereas among tumors greater than 45 mm in diameter, the proportion of Stage I cancers was 15% (N = 31,623), with intermediate size strata reflecting a declining proportion of Stage I lung cancers.[13] SEER and hospital registry data must be interpreted with care. These data reflect not only the underlying natural course of lung cancer but also the size-dependent probability of lung cancer detection.

Stage-shift refers to a shift in the stage distribution of lung cancers between two distinct populations. The presence of a stage-shift among small T1 lung tumors (<3 cm) is considered especially relevant to screening because screening has the ability to detect small lung cancers. In the SEER study, among tumors between 16 and 25 mm in diameter, 46% (N = 15,853) were Stage I, significantly smaller than the 54% proportion of Stage I cancers less than 15 mm in diameter.[13] Some studies of hospital registry data have confirmed this result[14,15] whereas others have failed to document this stage-shift.[16] Hospital registry data-sets may be up to two orders of magnitude smaller than SEER and biased toward surgical cases.[17] A stage-shift between chest X-ray detected and CT-detected incident cancers is expected because chest X-ray-detected cancers are larger, attributable to the lower sensitivity of chest X-ray.

Population-based models of lung cancer natural course have been previously developed to aid the interpretation of lung screening outcomes.[18,19] We introduce a novel estimation framework and specifically evaluate the consistency of estimates of radiologic size at cure threshold with both reported chest X-ray and CT outcomes.

Data, Methods, and Models

Mayo CT

The Mayo CT study was a prospective cohort study which began in 1999 and recruited 1,520 individuals (788 male, 732 female) greater than 50 years of age with a smoking history of more than 20 pack-years. Study participants received a baseline screen followed by four additional annual-repeat screens. We adopt the classification that prevalent lung cancers were visible at baseline, but possibly first detected during the study period. Incident cancers first became visible during the annual-repeat screening period. Interval cancers arose between annual-repeat screens. We obtained access to Mayo CT radiologic and pathologic records through participation in CISNET (www.cisnet.cancer.gov). Up to six nodules were reported for each study participant at the annual-repeat radiologic screen each year. We identified the largest among the reported nodules consistent with surgical location and timing of pathologic staging in order to reflect the radiologic size at detection among confirmed lung cancers. A total of 66 lung cancers were confirmed during the Mayo CT study.

MLP

The MLP was a RCT initiated in 1971 and completed on July 1, 1983 which recruited 10,933 male smokers over the age of 45 years. Of the original recruitment pool, 9,211 were accepted for randomization into a screening arm (N = 4,618) and a control arm (N = 4,593), based on a negative prevalence screen and the satisfaction of other entry criteria. MLP participants must have had an estimated life expectancy of five years or more, sufficient respiratory reserve to undergo pulmonary resection, no prior history of cancer of the lung or respiratory tract and no presentation of symptoms of lung cancer.[4] In the first seven trial years, there were 91 prevalence lung cancers, 151 incidence lung cancers (115 NSCLC) diagnosed in the screening arm and 120 incidence lung cancers (86 NSCLC) diagnosed in the control arm. We obtained the original radiology and pathology data reported during the MLP as well as extended mortality follow-up data.

Stage shift analysis

To simulate the age at lung cancer onset in the Mayo CT study, we adopt the two-stage clonal expansion (TSCE) model calibrated to the Nurses' Health Study and the Health Professionals Follow-Up Study.[20] The TSCE model incorporates gender and smoking-dependent parameter estimates

Table 1. Comparison of observed and expected non-small cell lung cancers in a simulation of the Mayo CT Study (males only)

	Simulated (CT Sensitivity = 0.9)	Simulated (CT Sensitivity = 0.8)	Observed
Prevalence (early-stage)	10.1	9.5	8
Prevalence (advanced-stage)	6.6	6.9	3
Incidence/Interval (early-stage)	11.2	10.9	7
Incidence/interval (advanced-stage)	1.9	2.1	7
Totals	**29.9**	**29.4**	**25**

Table 2. Radiologic sizes of non-small cell lung cancer incident and interval cancers in the Mayo Lung Project and the Mayo CT study (males only)

Mayo Lung Project		N (Known Size/Total)	Median Diameter (mm)	Mean Diameter (mm)
Screened Arm	Alive after 5 Years	43/49	25	26.35
	LC death within 5 years	43/55	35	36.44
Control Arm	Alive after 5 years	18/19	36.5	32.17
	LC death within 5 years	50/67	43.5	42.25
Mayo CT	No LC death	6/6	8.75	8.83
	LC death	7/8	12.5	15.07

and simulates the time of the appearance of the first malignant cell. After the appearance of the first malignant cell, we assume a fixed lag time of 6 months before the cancer becomes visible by CT. Given a model of tumor progression fit to the MLP, the lung cancer subsequently progresses through early and advanced stages governed by two independent exponential distributions.[5] Age and smoking-dependent other-cause mortality hazard functions are provided as a CISNET resource.

We simulate 10,000 iterations of the Mayo CT study and obtain expected counts of early and advanced-stage prevalent and incident NSCLCs (Table 1) among male participants. Higher sensitivity of CT relative to chest X-ray is expected to produce a stage-shift. We compute a test statistic based on the squared difference between expected and observed cancers in each of the four categories.

$$\mathrm{MSE} = \sum_{i=1}^{4} \frac{(\mathrm{Expected}_i - \mathrm{Observed}_i)^2}{\mathrm{Expected}_i}$$

The reported p-value is defined as the frequency of observing a MSE value as large or larger as the observed MSE statistic among the $N = 10{,}000$ simulations.

Probability of NSCLC death as a function of size at detection

For the 115 incident NSCLCs confirmed in the screened arm of the MLP, we collected data on radiologic tumor size and 5-year survival. Eleven NSCLCs were sputum-detected only, having unknown radiologic size and were omitted from the analysis. Among the 104 remaining NSCLCs, there were 52 lung cancer deaths, 11 other-cause deaths (9 Stage I, 2 Stage III) and four study-related deaths (3 Stage I, 1 Stage III) within 5 years of NSCLC detection. The remaining 37 individuals were alive after 5 years of NSCLC detection. For censored MLP cases, we impute Stage I cases as alive at five years and Stage II+ cases as resulting in lung cancer death. Among NSCLCs with reported radiologic sizes of either (*i*) not measurable but greater than 3 cm or (*ii*) not measurable but less than 3 cm, we impute the size as the median size among incidence NSCLCs in the similar size and survival category. Among the 14 interval and incidence NSCLCs in males detected in the Mayo CT, there were eight lung cancer deaths and six individuals were alive at the end of the follow-up period. The median follow-up time after initial detection for survivors was 4.0 years. In all cases, we summarize the tumor diameter as the average of the reported long and short axes. A comparison of the reported tumor diameter information between the Mayo CT and MLP is given in Table 2.

We assume the conditional cumulative density function of LC death as a function of tumor diameter at detection follows a gamma distribution with shape and scale parameters, alpha (α) and gamma (γ), respectively. Under the minimal assumption that tumors increase uniformly in size, we define a likelihood function as follows:

For detected NSCLCs with diameter s_0, resulting in no LC death:

$$\Pr\left[\mathrm{LC\ Death} = 0 \,|\, \mathrm{Detection} = s_0\right] = \Pr\left[X > s_0 \,|\, \alpha, \gamma\right]$$

For detected NSCLCs with diameter s_1, resulting in a LC death:

$$\Pr\left[\mathrm{LC\ Death} = 1 \,|\, \mathrm{Detection} = s_1\right] = \Pr\left[X < s_1 \,|\, \alpha, \gamma\right]$$

Table 3. Median tumor diameter (mm) at cure threshold as a function of doubling time (Ln (2)/*f*) and Log_{10} (mutation rate (μ))

	Log_{10} (mutation rate (μ))									
Doubling Time (Days)	−5.15	−4.95	−4.75	−4.55	−4.35	−4.15	−3.95	−3.75	−3.55	−3.35
40	100.3	100.3	100.2	100.2	77.4	46.5	28.3	16.9	10.2	6.5
80	100.1	100.1	100.1	71.5	44.7	26.8	16.6	10.1	6.5	5.0
120	100.1	100.1	91.0	53.8	30.2	19.8	12.7	7.8	5.1	5.0
160	100.1	100.0	73.2	44.2	26.9	16.4	10.1	6.7	5.0	5.0
200	100.0	100.0	60.5	37.8	21.3	13.6	8.8	5.7	5.0	5.0
240	100.0	92.4	53.6	32.2	20.2	12.5	7.9	5.0	5.0	5.0
280	100.0	80.5	47.0	29.2	18.3	11.3	7.0	5.0	5.0	5.0

Asymptotic properties of maximum likelihood statistical methods are well defined.[21] Maximum likelihood estimates (MLEs) for the shape parameter, α, and the scale parameter, γ, are obtained under a bounded variance constraint: variance is assumed bounded by the sample variance of radiologic sizes at detection in the pooled group. We estimate the MLE vector (x) for the pooled sample of Mayo CT and MLP lung cancers as well as for each individual data-set separately. Based on x and the resulting Hessian matrix, Σ^{-1}, we report a 95% confidence interval for the median of the distribution for each data-set independently. The 95% confidence interval for the median is computed on parameter values of x falling within the 95% confidence ellipse given by: $x \in (x - \mu)'\Sigma^{-1}(x - \mu) < \chi^2_{2,0.95}$. We conduct a likelihood ratio test to examine the strength of the evidence that a unique parameter vector is consistent with data from both the MLP and Mayo CT. We compute:

$$-2\ \ln\frac{L(\hat{\alpha},\hat{\gamma}|s)_{\mathrm{MLP}}L(\hat{\alpha},\hat{\gamma}|s)_{\mathrm{Mayo}}\mathrm{CT}}{L(\hat{\alpha},\hat{\gamma}|s)_{\mathrm{pooled}}}$$

where *s* is the vector of detection sizes and obtain the resulting *p*-value based on a chi-squared distribution having two degrees of freedom. We compute the probability that a 10-mm NSCLC detected by chest X-ray and CT will result in a lung cancer death with the associated 95% confidence interval.

Multi-type branching process model

We describe a multi-type branching process model in which cancer arises from a single stem-like malignant progenitor cell. Upon cell division, this progenitor cell gives rise to two identical stem-like cells with probability *f* and gives rise to one stem-like progenitor cell and one terminally differentiated cell with probability 1-*f*. Tumor growth occurs as the result of the accumulation of cells from multiple cell divisions. We assume a pure-birth process in which terminally differentiated cells contribute to the overall size of the tumor but, unlike stem-like progenitor cells, do not undergo further cell division events. The accumulation of *n* key mutations in a stem-like progenitor cell results in the development of an incurable tumor, ultimately resulting in a lung cancer death. We define the tumor diameter at which *n* mutations have accumulated as the cure threshold. The probability of any of the *n* key mutations occurring at a cell division step equals μ.

We assume the average volume occupied by a single cell of clonal origin is 50 μm. Conversion between cell number and tumor diameter is based on the volume equation $V = 4/3\ \pi r^3$. We allow the range of parameter values to encompass a spectrum of doubling times ranging from 20 to 300 days and a stem cell fraction of up to 5%. The median size at the cure threshold over this parameter range for *f* and μ is represented in Table 3.

We collect data on tumor size, survival, and screen history for incident NSCLCs arising during the MLP and Mayo CT studies. We simulate a joint uniform vector [U_1, U_2] representing VDTs [range: 20:300] and median size at cure threshold [range: 5:100]. We determine the values of *f* and μ corresponding to [U_1, U_2] by the relationship $U_1 = \ln(2)/f$ and by inverse mapping of an expanded Table 3. We generate the resulting tumor trajectory and retain the values of *f* and μ, contingent on the following:

1. For NSCLCs resulting in no lung cancer death (curable), the simulated size at cure threshold is larger than the size at detection.
2. For NSCLCs resulting in a lung cancer death (incurable), the simulated size at cure threshold is smaller than the size at detection.
3. Given a screen history before detection (*f*, μ) if and only if $U > 1 - e^{-\alpha d^{\beta}}$ for each tumor size at the timing of each of *p* prior screens, where *U* is a uniform random variable on [0,1]. We assume a Weibull screen detection probability function having a 50% probability of detection for a 30-mm nodule.

We assume a NSCLC is visible at the earliest screen among repeat screens within a window of 75 days directly before lung cancer diagnosis. Chest X-ray screens may be performed after nodule detection but before lung cancer diagnosis as part of the diagnostic work-up. For every

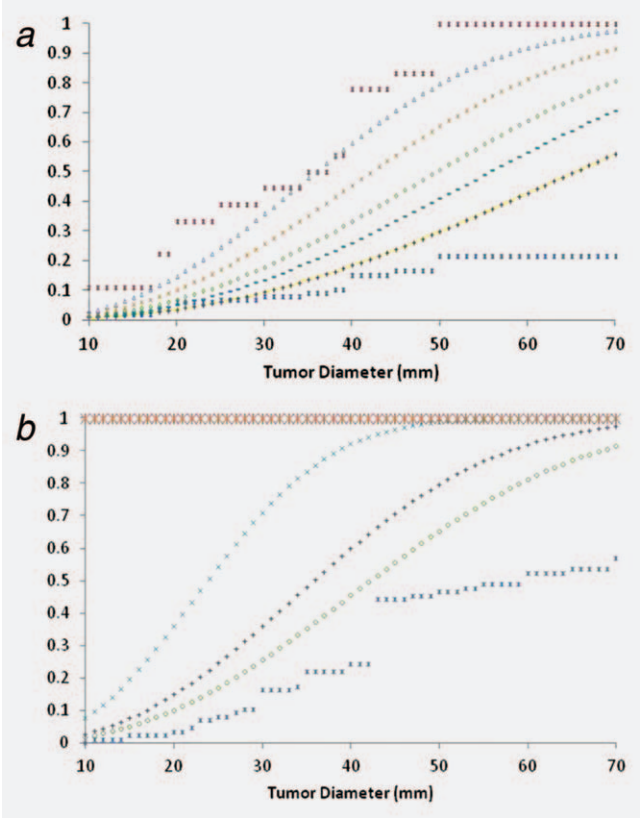

Figure 1. Cumulative density functions (cdfs) representing the probability of detection as a function of tumor diameter for (*a*) curable lung cancers and (*b*) incurable lung cancers in the absence of screening. (*a*) Kaplan-Meier upper and lower bounds and intermediate Weibull (α,β) cdfs representing the probability of detection of curable cancers in the absence of screening (1). (*i*) Upper and lower Kaplan-Meier bounds (*), Weibull (9E-05,2.5) (Δ), Weibull (6E-05,2.5) (x), Weibull (4E-05,2.5) (o), Weibull (3E-05,2.5) (---), Weibull (2E-05,2.5) (+). b) Kaplan-Meier upper and lower bounds and intermediate Weibull (α,β) cdfs representing the probability of detection of incurable cancers in the absence of screening (2). (*ii*) Upper and lower Kaplan-Meier bounds (*), Weibull (2.5E-02,2.5) (*), Weibull (2.5E-04,2.5) (x), Weibull (9E-05,2.5) (+), Weibull (6E-05,2.5) (o). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

screened arm NSCLC having known tumor diameter at detection, we continue replacement sampling until 1,000 values of f and μ are generated.

We estimate size-dependent clinical detection in the control arm using Kaplan-Meier (KM) estimation, separately for curable and incurable NSCLCs. We estimate upper and lower KM bounds, based on minimum and maximum numbers of NSCLCs at risk of detection for each size, respectively. The true number of NSCLCs at risk of curable or incurable detection for each tumor size is unknown. For example, it is not known whether a 40-mm incurable NSCLC had been incurable at 20 mm. We identify a set of Weibull distribution functions that fall within the upper and lower KM bounds to represent the size-dependent detection probability of control arm cancers (Fig. 1).

For each parameter pair (f, μ), we compute the probability that a NSCLC is detected in the control arm. We assume a uniform distribution governing the onset of detectable NSCLC within the screening window and specify the Weibull probability functions governing NSCLC detection in the control arm. We consider only trajectories of (f, μ) for which the NSCLC fails to be detected at the prevalence screen but would be screen-detectable by a single screen at the end of the follow-up period. The precision of the probability estimates for each parameter pair (f, μ) is based on 10,000 simulated trajectories. Based on the estimation principle of method of moments (MOM), we define a solution as a set of parameter values (f_i, μ_i) satisfying:

$$\Sigma_{i=1}^{N} p_1(f_i, \mu_i) = C_1 \quad \text{and} \quad \Sigma_{i=1}^{N} p_2(f_i, \mu_i) = C_2$$

where p_1 = probability of curable NSCLC detection, p_2 = probability of incurable NSCLC detection, N = the number of screened arm-detected NSCLCs, C_1 = the number of curable control arm NSCLCs, and C_2 = the number of incurable control arm NSCLCs. We opt for a MOM estimation approach to eliminate model constraints associated with maximum likelihood estimation. To obtain a MOM solution to the parameter space, we first draw a random set of N screened-arm NSCLCs. If a selected screened-arm NSCLC has missing tumor size information, we select a set of parameter values consistent with the NSCLC having the identical survival status but known size information. We allow C_1 and C_2 to vary according to a multinomial distribution with proportions of curable, incurable and missed NSCLCs equal to C_1/N, C_2/N, and $(N\text{-}C_1\text{-}C_2)/N$, respectively. We search the parameter space to obtain up to 1,000 solutions for each of the 20 detection scenarios.

For each of the 20 detection scenarios, we report the median VDTs and median tumor diameters at cure threshold among all screened-arm detected NSCLCs and for the subset of incurable NSCLCs detected in the control arm, averaged over 100 MOM-calibrated models. We also report the average median tumor diameters at detection for curable and incurable NSCLCs and the proportion of curable NSCLCs in the control arm of the MLP (Table 4). We select four representative MOM-calibrated models from different detection scenarios. To simulate screened-arm outcomes in the MLP, we assume a screen detection function having 99% probability of detection for a 30-mm nodule with 75% of NSCLCs in a detectable location of the chest. We simulate both arms of the MLP and the single arm of the Mayo CT study and report the simulation outcomes reflecting goodness of each model fit (Table 5).

Results

In the Mayo CT, the proportion of lung cancer deaths among incident NSCLCs (8/14) in males was similar to the proportion of lung cancer deaths among incident NSCLCs detected in the MLP (55/104). Our existing simulation model predicts a stage-shift among incident NSCLCs relative to the MLP, resulting in a MSE p-value of either

Table 4. Mayo lung project simulation outcomes conditional on models of lung cancer progression calibrated using method of moments estimation for specified detection scenarios

				Mayo lung project control arm				
Detection scenario	Weibull parameters (α,β)[1]	Median diameter at cure threshold (all screen) (mm)	Volume doubling time (all screen) (days)	Median diameter at cure threshold (incurable only) (mm)	Volume doubling time (incurable) (days)	Median diameter at curable detection (mm)	Median diameter at incurable detection (mm)	Percent Cure
Observed	N/A	N/A	N/A	N/A	N/A	36.5	46.5	22.1%
1	(9E-05,2.5) (6E-05,2.5)	14.4	85.8	11.4	75.7	22.7	41.8	20.7%
2	(9E-05,2.5) (9E-05,2.5)	14.6	99.4	11.5	87.3	23.2	36.1	21.6%
3	(9E-05,2.5) (2.5E-04,2.5)	14.7	120.7	11.3	106.1	23.5	25.6	21.8%
4	(9E-05,2.5) (2.5E-02,2.5)							
5	(6E-05,2.5) (6E-05,2.5)	17.2	83.0	13.6	74.0	26.6	42.4	20.6%
6	(6E-05,2.5) (9E-05,2.5)	16.8	97.1	13.2	85.9	27.1	36.8	20.9%
7	(6E-05,2.5) (2.5E-04,2.5)	17.5	117.5	13.1	103.3	27.7	26.6	21.7%
8	(6E-05,2.5) (2.5E-02,2.5)							
9	(4E-05,2.5) (6E-05,2.5)	20.5	82.6	16.2	74.1	31.4	43.5	20.6%
10	(4E-05,2.5) (9E-05,2.5)	20.9	90.3	15.9	81.0	32.1	38.1	21.5%
11	(4E-05,2.5) (2.5E-04,2.5)	20.6	117.4	15.2	105.0	32.8	27.8	21.7%
12	(4E-05,2.5) (2.5E-02,2.5)	24.4	142.1	16.1	124.4	34.0	16.1	24.5%
13	(3E-05,2.5) (6E-05,2.5)	23.3	79.7	18.2	72.0	35.1	44.4	20.5%
14	(3E-05,2.5) (9E-05,2.5)	23.2	86.2	17.8	77.5	35.3	39.1	20.5%
15	(3E-05,2.5) (2.5E-04,2.5)	23.8	110.9	17.5	98.9	36.5	29.0	21.6%
16	(3E-05,2.5) (2.5E-02,2.5)	26.5	134.4	17.9	119.3	37.4	17.9	23.4%
17	(2E-05,2.5) (6E-05,2.5)	28.3	74.9	22.1	68.7	40.8	46.3	20.5%
18	(2E-05,2.5) (9E-05,2.5)	28.2	80.1	21.7	72.9	41.2	41.0	20.5%
19	(2E-05,2.5) (2.5E-04,2.5)	29.1	96.3	21.4	86.4	42.2	31.2	21.5%
20	(2E-05,2.5) (2.5E-02,2.5)	31.6	128.2	21.3	114.4	42.7	21.3	22.6%

[1]Weibull parameters for clinical detection for curable and incurable cancers, respectively.

Table 5. Mayo lung project and Mayo CT simulation outcomes for selected individual tumor progression models obtained by method of moments estimation

	Mayo lung project						Mayo CT		
	Screened arm			Control arm			Screened arm		
	Observed	Simulated	90% Range	Observed	Simulated	90% Range	Observed	Simulated	90% Range
Detection Scenario #2 MDCT = 18.2/13.1 mm VDT = 136/129 days									
Curable lung cancers	49	42	(32, 52)	19	23	(15, 32)	6	8	(4, 13)
Incurable lung cancers	55	51	(40, 63)	67	61	(51, 75)	8	5	(2, 8)
Total lung cancers	104	94	(78, 108)	86	84	(72, 100)	14	13	(8, 19)
Median diameter, curable lung cancers	25	14.9	(13.0, 17.4)	36.5	23.9	(18.8, 29.6)	8.75	7.7	(6.5, 9.6)
Median diameter, incurable lung cancers	35	27.3	(23.0, 31.8)	43.5	36.3	(32.7, 39.6)	12.5	26.4	(11.57, 76)
Detection Scenario #9 MDCT = 21.5/16.3 mm VDT = 100/85 days									
Curable lung cancers	49	41	(31, 51)	19	20	(13, 28)	6	8	(4, 13)
Incurable lung cancers	55	54	(42, 67)	67	68	(55, 80)	8	5	(2, 9)
Total lung cancers	104	95	(79, 110)	86	88	(73, 103)	14	14	(8, 19)
Median diameter, curable lung cancers	25	16.9	(14.4, 19.8)	36.5	32.7	(25.3, 40.1)	8.75	8.31	(6.8, 10.43)
Median diameter, incurable lung cancers	35	33.0	(27.4, 39.1)	43.5	43.8	(39.8, 47.9)	12.5	26.1	(12.9,67)
Detection Scenario #14 MDCT = 26.5/18.5 mm VDT = 94/89 days									
Curable Lung Cancers	49	45	(35, 56)	19	20	(13,29)	6	9	(4, 14)
Incurable lung cancers	55	50	(38, 61)	67	66	(54,80)	8	5	(2, 9)
Total lung cancers	104	95	(78, 111)	86	86	(71,102)	14	14	(8, 21)
Median Diameter, curable lung cancers	25	17.0	(14.7,20.2)	36.5	37.2	(28.5, 45.9)	8.75	8.4	(6.8, 11.8)
Median diameter, incurable lung cancers	35	31	(26.6, 36.9)	43.5	39.3	(35.9,43.1)	12.5	25.8	(11.7, 68.6)
Detection Scenario #18 MDCT = 31.7/24.8 mm VDT = 99.5/87 days									
Curable lung cancers	49	50	(38,62)	19	21	(14,28)	6	9	(4, 14)
Incurable lung cancers	55	45	(36,57)	67	68	(55,80)	8	4	(1, 8)
Total lung cancers	104	95	(78,111)	86	89	(75,102)	14	13	(7, 20)
Median diameter, curable lung cancers	25	17.6	(15,20.5)	36.5	41.9	(32.2,51.5)	8.75	8.8	(7.1, 11.8)
Median diameter, incurable lung cancers	35	35.5	(30.6,40.8)	43.5	41.5	(38.2,45.6)	12.5	34.0	(12.1, 89.7)

MDCT = median diameter at cure threshold (all screen-detected/incurable cancers only).
VDT = volume doubling times (all screen-detected/incurable cancers only).

0.0026 or 0.0077, given CT detection sensitivity of either 90 or 80%, respectively.

Among the 104 incident NSCLCs detected in the screened-arm of the MLP, there were 41 squamous-cell lung cancers, 39 adenocarcinomas, 23 large-cell lung cancers, and one cancer of another histologic sub-type. Association between 5-year NSCLC survival and histologic sub-type was not detected by chi-square association analysis (p-value= 0.6529). ANOVA analysis failed to detect a difference among the mean NSCLC tumor diameters at detection for any histologic subtype for either curable (p-value = 0.6503) or incurable NSCLCs (p-value = 0.7893). Among the 14 incident NSCLCs detected in the Mayo CT, there were seven squamous-cell lung cancers, two adenocarcinomas, and five other histologic sub-types. With the exception of one incurable Stage IIB squamous-cell lung cancer with a tumor diameter of 33 mm, all incident NSCLCs in the Mayo CT study among males had a tumor diameter less than 20 mm.

Based on a likelihood ratio test, the probability of 5-year NSCLC survival as a function of tumor diameter differs significantly by detection modality (p = 0.0312). In the MLP, the estimated tumor diameter at detection at which 50% of NSCLCs are curable is 27.61 mm [95% CI: (17.7, 33.3)] whereas in the Mayo CT, the estimated tumor diameter at detection for which 50% of NSCLCs are curable is 10.14 mm [95% CI: (0.13, 13.8)]. In the MLP, the estimated probability of lung cancer death given a 10 mm detected NSCLC is 7.1%[95% CI: (3.4%,27.8%)] whereas in the Mayo CT, the estimated probability of lung cancer death given a 10 mm detected NSCLC is 48.5% [95% CI: (17.8%,99%)].

We successfully identified parameter sets over all but two of the 20 detection scenarios satisfying the MOM estimation criteria. The median tumor diameter at cure threshold averaged over 100 MOM parameter sets varies by detection scenario and ranges from 14.4 to 31.6 mm. In a simulation of the MLP, the close approximation to the observed proportion of 22% curable NSCLCs among control-arm cancers reflects the MOM estimation criteria. The estimates of median tumor diameter at cure threshold among the subset of incurable detected control-arm NSCLCs in the MLP varied by detection scenario and ranged in our simulations from 11.3 mm to 22.1 mm.

We incorporated four specific MOM parameter sets into the tumor progression component of the simulation models of the MLP and Mayo CT studies (Table 5). For each of the simulations, the number of curable and incurable incident NSCLCs detected in the Mayo CT and their respective tumor diameters at detection lie within or nearly within the 90% range of simulated outcomes. In the MLP simulation, the total number of NSCLCs in the control arm is closely predicted, as is the observed proportion of incurable lung cancers (22%). In all simulations excepting detection scenario two, the median tumor diameters at detection lie within or nearly within the 90% range of simulated outcomes for control-arm NSCLCs. With respect to screened-arm MLP outcomes, an exception to the goodness of fit is the under-estimation of the median tumor diameters at detection of curable NSCLCs.

Discussion

An existing simulation model of lung cancer natural course predicts a stage-shift among incident lung cancers in the Mayo CT relative to the MLP. Simulations proved incompatible with the observed Mayo CT incident data despite the smaller sizes of Mayo CT incident lung cancers. These results appear to contradict SEER data demonstrating a decreasing proportion of Stage I lung cancers with increasing tumor diameter. However, in both the Mayo CT and MLP individually, early-stage lung cancers were smaller than advanced-stage cancers, consistent with SEER trends. Between data-sets, advanced-stage cancers in the Mayo CT were smaller than early-stage cancers detected in the MLP.

CT is able to detect lung cancers at small sizes not previously seen by usual care nor by chest X-ray. An advanced, incurable lung cancer which may have been detected by usual care at a size of 45 mm may now be detected by CT having a tumor diameter of 15 mm. This lung cancer, not previously observed at 15 mm, may still be an advanced, incurable lung cancer. We find that detection modality is a significant factor in the prediction of 5-year lung cancer survival as a function of tumor size. The broad implication of this finding is that small tumors detected by CT may be more aggressive than small lung cancers seen previously.

Our estimates of the median tumor diameter at cure threshold are substantially smaller than median diameters of advanced-stage lung tumors in SEER. It has been noted that bias in SEER in the form of underrepresentation of small, asymptomatic early-stage lung cancers and under-reported size information among large, unresected advanced-stage lung cancers is likely to strengthen the reported relationship between tumor size and stage.[13] However, under-representation in SEER of small, asymptomatic advanced-stage cancers could weaken the reported relationship between tumor size and stage. Data from the lung screening study, the NLST pilot study, demonstrated a two-fold higher detection rate for advanced-stage lung cancers in the CT screened arm compared to the chest X-ray screened arm, thereby challenging the assumption that advanced-stage cancers remain asymptomatic only for a negligible period of time.[22]

Our cure threshold models are fit to data from the MLP, for which lung cancers were detected and treated between 1971 and 1983. If 5-year survival trends have significantly improved over the past 30 years, then the cure threshold models we produce may be anachronistic when applied to today's screening methods. Five-year lung cancer survival remains poor. According to a study of SEER survival trends, the 5-year NSCLC survival among males was 13.4% over the period 1978 to 1980, improving to 15.4% between 1999 to 2005,[23] a modest 2% improvement in NSCLC survival over a

period of 25 years. Furthermore, the PLCO screening study confirms the nearly 30-year old MLP results reporting no mortality benefit associated with early detection by chest X-ray.[2]

The NLST has been called the "best umpire in town," in determining whether a lung cancer mortality reduction attributable to CT exists.[24] However, variation in CT screening outcomes across studies is notable. The frequency of Stage I annual-repeat lung cancers was 48.6% (17/35) in the Mayo CT study, 46.8% (15/32) in the DANTE CT screening study, similar to the 48.6% (68/140) Stage I annual-repeat cancers observed in the first 7 years of the MLP, when sputum-detected cases are excluded.[4,25,26] The New York ELCAP reported that among 48 lung cancers detected in annual repeat screens and not visible at baseline, 79.2% (38/48) were classified as N0M0.[27] In the Dutch NELSON study, 73.7% (42/57) cancerous nodules detected in the second round of screening were classified as pathological stage I.[28]

Variation in CT mortality outcomes presupposes the observed variation in stage distribution among CT-detected incident cancers. In a matched cohort analysis, estimates of mortality reduction in the New York component of the I-ELCAP screening study are 36% and 64% when data are compared to the CPS-II and CARET studies, respectively.[29] A distinguishing feature of the I-ELCAP protocol is the aggressive management of annual-repeat lung cancers: biopsy by fine-needle aspiration is recommended if growth is seen for a new non-calcified nodule larger than 5 mm, as measured at a repeat CT scan after one month of detection.[30] If a large proportion of aggressive NSCLCs have a cure threshold below 20 mm, a high sensitivity to variation among CT screening protocols is expected.

It has been suggested that increasing the size threshold for biopsy will reduce the false positive rate among CT-detected lung cancers and reduce overdiagnosis.[31] The effect of increasing the size threshold on the detection of true positive cases at an early-stage is presumed small, but this depends on a comprehensive understanding of lung cancer natural course. If raising the size threshold for biopsy leads to many NSCLCs being missed at a curable size, then the ratio of false positive cases to curable true positive cases may, in fact, increase.

In conclusion, the NLST study has demonstrated that lung cancer mortality can be reduced by CT screening. Questions remain concerning the optimal implementation of early detection strategies.

References

1. Aberle DR, Adams AM, Berg CD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011;365: 395–409.
2. Oken MM, Hocking WG, Kvale PA, et al. Screening by chest radiograph and lung cancer mortality: the prostate, lung, colorectal, and ovarian (PLCO) randomized trial. *JAMA* 2011; 306:1865–73.
3. Marcus PM, Bergstralh EJ, Fagerstrom RM, et al. Lung cancer mortality in the Mayo lung project: impact of extended follow-up. *J Natl Cancer Inst* 2000;92:1308–16.
4. Fontana RS, Sanderson DR, Woolner LB, et al. Lung cancer screening: the Mayo Program. *J Occup Med* 1986;28:746–50.
5. Flehinger BJ, Kimmel M, Polyak T, et al. Screening for lung cancer: the Mayo lung project revisited. *Cancer* 1993;72:1573–80.
6. Marcus PM, Bergstralh EJ, Zweig MH, et al. Extended lung cancer incidence follow-up in the Mayo lung project and overdiagnosis. *J Natl Cancer Inst* 2006;98:748–56.
7. Goldwasser DL, Kimmel M. Modeling excess lung cancer risk among screened arm participants in the Mayo lung project. *Cancer* 2010;116: 122–31.
8. Dominioni L, Poli A, Mantovani W, et al. Volunteer effect and compromised randomization in the Mayo lung project of screening for lung cancer. *Eur J Epidemiol* 2011;26:79–80.
9. Pashkevich MA, Sigal BM, Plevritis SK. Modeling the transition of lung cancer from early to advanced stage. *Cancer Causes Control* 2009;20: 1559–69.
10. Detterbeck FC, Gibson CJ. Turning gray: the natural history of lung cancer over time. *J Thorac Oncol* 2008;3:781–92.
11. Lindell RM, Hartman TE, Swensen SJ, et al. Five-year lung cancer screening experience: CT appearance, growth rate, location, and histologic features of 61 lung cancers. *Radiology* 2007;242: 555–62.
12. Yankelevitz DF. Quantifying "overdiagnosis" in lung cancer screening. *Radiology* 2008;246:332.
13. Wisnivesky JP, Yankelevitz D, Henschke CI. Stage of lung cancer in relation to its size: Part 2. Evidence. *Chest* 2005;127:1136–39.
14. Yang F, Chen H, Xiang J, et al. Relationship between tumor size and disease stage in non-small cell lung cancer. *BMC Cancer* 2010;10:474.
15. Flieder DB, Port JL, Korst RJ, et al. Tumor size is a determinant of stage distribution in T1 non-small cell lung cancer. *Chest* 2005;128:2304–08.
16. Heyneman LE, Herndon JE, Goodman PC, et al. Stage distribution in patients with a small ($\leq$3 cm) primary nonsmall cell lung carcinoma: implication for lung carcinoma screening. *Cancer* 2001;92:3051–55.
17. Yankelevitz D, Wisnivesky JP, Henschke CI. Stage of lung cancer in relation to its size: Part 1. Insights. *Chest* 2005;127:1132–35.
18. Flehinger BJ, Kimmel M. The natural history of lung cancer in a periodically screened population. *Biometrics* 1987;43:127–44.
19. McMahon PM, Kong CY, Johnson BE, et al. Estimating long-term effectiveness of lung cancer screening in the Mayo CT screening study. *Radiology* 2007;248:278–87.
20. Meza R, Hazelton WD, Colditz GA, et al. Analysis of lung cancer incidence in the nurses' health and the health professionals' follow-up studies using a multistage carcinogenesis model. *Cancer Causes Control* 2008;19:317–28.
21. Mardia KV, Kent JT, Bibby JM. Multivariate analysis. San Diego: Academic Press; 1979.
22. Gohagan JK, Marcus PM, Fagerstrom RM, et al. Final results of the lung screening study, a randomized feasibility study of spiral CT versus chest X-ray screening for lung cancer. *Lung Cancer* 2005;47:9–15.
23. American Lung Association Epidemiology and Statistics Unit. Trends in Lung Cancer Morbidity and Mortality 2010. Available at: http://www.lung.org/finding-cures/our-research/trend-reports/lc-trend-report.pdf.
24. Swensen SJ, Jett JR, Midthun DE, et al. Computed tomographic screening for lung cancer: home run or foul ball? *Mayo Clin Proc* 2003;78:1187–88.
25. Swensen SJ, Jett JR, Hartman TE, et al. CT screening for lung cancer: five-year prospective experience. *Radiology* 2005;1:259–65.
26. Infante M, Cavuto S, Lutman F, et al. A randomized study of lung cancer screening with spiral computed tomography: three year results from the DANTE trial. *Am J Respir Crit Care Med* 2009;180:445–53.
27. Carter D, Vazquez M, Flieder DB, et al. Comparison of pathologic findings of baseline and annual repeat cancers diagnosed on CT screening. *Lung Cancer* 2007;56:193–99.
28. Van Klaveren RJ, Oudkerk M, Prokop M, et al. Management of lung nodules detected by volume CT scanning. *N Engl J Med* 2009;361: 2221–29.
29. Henschke CI, Boffetta P, Gorlova O, et al. Assessment of lung cancer mortality reduction from CT screening. *Lung Cancer* 2010;71:328–32.
30. Henschke CI, Yankelevitz DF, Libby DM, et al. Survival of patients with stage I lung cancer detected on CT screening. *N Engl J Med* 2006; 1763–71.
31. Esserman L, Thompson I. Solving the overdiagnosis dilemma. *J Natl Cancer Inst* 2010; 102:582–83.