Supplementary Online Content

Patz EF, Pinsky P, Gatsonis C, Sicks JD, Kramer BS, Tammemägi MC, Chiles C, Black WC, Aberle DR; for the NLST Overdiagnosis Manuscript Writing Team. Overdiagnosis in low-dose computed tomography screening for lung cancer. *JAMA Intern Med*. Published online December 9, 2013.
doi:10.1001/jamainternmed.2013.12738

eAppendix. Convolution Model

eTable. Model Fit to NLST Data—Counts of Lung Cancers

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

eAppendix. Convolution Model

The convolution model defines a pre-clinical phase (before symptoms), a sojourn time distribution, and a sensitivity of screening [10-11]. Sojourn time is the period spent in the pre-clinical phase, before the disease becomes clinically apparent. A cubic model was utilized to fit pre-clinical incidence as a function of age. An exponential distribution was utilized for sojourn time; this implies that mean lead time equals mean sojourn time. Pre-clinical incidence and sojourn time were assumed independent of trial arm. Note that pre-clinical incidence, or initiation into the pre-clinical phase occurs when the tumor is first theoretically detectable; for the current study this would be an approximately 4 mm (cancerous) pulmonary nodule. Separate sensitivity parameters were fit for LDCT and CXR. The model parameters were fit using maximum likelihood [10].

The likelihood ratio test showed that fitting separate models for BAC and non-BAC (with 6 extra parameters) resulted in a significantly greater overall likelihood ($p < 0.001$) than fitting one model for all NSCLC. In fitting the model, it was assumed that there was no screening outside of the NLST protocol; for example, all cancers diagnosed after the screening period in NLST (T3-T7) were assumed not to be screen-detected. Confidence intervals on parameters were obtained through the profile likelihood method.

Utilizing the fitted model parameters from above, we ran a simulated cohort of 500,000 individuals undergoing a given screening and follow-up regimen (e.g., 5 annual screens and 10 years of total follow-up) to estimate the numbers of screen-detected and total cancers diagnosed under LDCT, CXR and no screening. To approximate the NLST scenario, the cohort was assumed to start screening at age 60. Denote N_{LDCT}, N_{CXR} and N_{None} the total number of diagnosed cases under LDCT screening, CXR screening and no screening, respectively, and $N_{LDCT,SC}$ be the number of LDCT screen-detected cases. Then $P_{S,None}$=(

$N_{LDCT} - N_{None})/ N_{LDCT,SC}$ and $P_{S,CXR}=(N_{LDCT} - N_{CXR})/ N_{LDCT,SC}$ where the former is excess cancers (with LDCT) relative to no screening and the latter is relative to CXR screening. Confidence intervals on excess cancer rates were computed by bootstrapping. Specifically, bootstrapping was used to generate a sample of estimated parameter vectors and excess cancer rates were then calculated from these parameter vectors using simulations.

eTable. Model Fit to NLST Data—Counts of Lung Cancers

		LDCT Arm		CXR Arm	
		Screen Detected	Interval	Screen Detected	Interval
Subset	Study Year	OBS|EXP	OBS|EXP	OBS|EXP	OBS|EXP
NSCLC Excluding BAC	T0	213|244	14|25	113|96	36|61
	T1	124|108	11|20	56|76	48|54
	T2	159|110	15|20	65|77	44|52
	Post Screening [1]		269|280		378|338
	Never Screened		10|5		17|9
	All	496|462	319|350	234|249	523|514
BAC	Study Year	OBS|EXP	OBS|EXP	OBS|EXP	OBS|EXP
	T0	38|39	0|2	8|4	0|3
	T1-T2	57|54	1|3	5|9	4|7
	Post-Screening [1]		14|8		18|15
	Never Screened		1|0		1|0
	All	95|93	16|13	13|13	23|25
Total	Total NSCLC	OBS|EXP= 926|918		OBS|EXP= 793|801	

[1]T3 and beyond

Note: EXP is expected under model fit. OBS is observed in NLST.