



Published in final edited form as:

Ann Intern Med. 2018 July 03; 169(1): 1–9. doi:10.7326/M17-2561.

Identifying Patients for Whom Lung Cancer Screening is Preference-Sensitive: A Microsimulation Study

Tanner J. Caverly, MD, MPH^{1,2}, Pianpian Cao, MPH³, Rodney A. Hayward, MD^{1,2}, and Rafael Meza, PhD^{2,3}

¹VA Center for Clinical Management Research, Ann Arbor and University of Michigan Medical School

²Institute for Health Policy Innovation, University of Michigan, Ann Arbor, Michigan

³School of Public Health, University of Michigan, Ann Arbor, Michigan

Abstract

Background—Many health systems are still exploring how to implement an effective, patient-centered low-dose computed tomography (LDCT) screening program.

Objective—Examine factors that influence when LDCT screening is preference-sensitive.

Design—State-transition microsimulation model

Data Sources—Two large randomized trials, published decision analyses, and the Surveillance, Epidemiology and End-Results cancer registry

Target Population—US-representative sample of simulated patients meeting current US Preventive Services Task Force screening eligibility criteria

Time Horizon—Lifetime

Perspective—Individual

Intervention—LDCT screening annually for 3 years

Corresponding Author: Tanner J. Caverly, MD, MPH, VA Center for Clinical Management Research and University of Michigan Medical School, 2800 Plymouth Rd, Building 16, Rm 331W, Ann Arbor, MI 48109 tcaverly@med.umich.edu Phone: 734-222-8958.

Current Author Addresses:

Drs. Caverly, Hayward and Meza: Institute for Health Policy Innovation, University of Michigan, Ann Arbor, Michigan

Dr. Caverly: VA Center for Clinical Management Research, Ann Arbor and University of Michigan Medical School, North Campus Research Complex Building 16, Rm 321W, Ann Arbor, MI 48105

Ms. Cao: School of Public Health, University of Michigan, M5336 SPH II, 1415 Washington Heights, Ann Arbor, Michigan 48109-2029

Dr. Hayward: VA Center for Clinical Management Research, Ann Arbor and University of Michigan Medical School, 2800 Plymouth Rd., Bldg. 10, Rm. G016-4A, Ann Arbor, Michigan 48109-2800

Dr. Meza: School of Public Health, University of Michigan, M5533 SPH II, 1415 Washington Heights, Ann Arbor, Michigan 48109-2029

Primary Funding Source

none

Registration: n/a

IRB approval: Exempt

Reproducible Research Statement: *Study protocol and statistical code:* Not available. *Data set:* Available from the National Cancer Institute Cancer Data Access System (www.biometry.nci.nih.gov/cdas/nlst).

Outcomes Measures—Lifetime quality-adjusted life-year gains and reduction in lung cancer mortality. To examine the effect of preferences on net benefit, we varied disutilities (i.e., negative feelings) quantifying the burden of screening and follow-up across a likely range. We also examined the effect of varying the rate of false-positive scans and overdiagnosis associated with screening.

Results of Base-Case Analysis—Moderate differences in preferences about the downsides of LDCT screening influenced whether screening was appropriate for eligible persons with < 0.3% annual lung cancer risk or life-expectancy < 10.5 years. For higher-risk eligible persons with longer life-expectancy, roughly 50% of the study population, LDCT screening overcame even highly negative views about screening and its downsides.

Results of Sensitivity Analysis—Rates of false-positive findings and overdiagnosed lung cancers were not highly influential.

Limitation—The quantitative thresholds we identified may vary depending on the structure of the microsimulation model.

Conclusions—Identifying circumstances under which LDCT screening is more vs. less preference-sensitive may help clinicians personalize their approach to discussing LDCT screening, tailoring to both preferences and clinical benefit.

Introduction

In 2011, the National Lung Screening Trial (NLST) found that annual lung cancer screening with low-dose computed tomography (LDCT screening) can substantially reduce mortality from lung cancer (1), the leading cause of cancer death in the US (2). Many health systems are still exploring how to best implement an LDCT screening program (3).

Several factors complicate LDCT implementation: 1) the absolute risk reduction with screening varies greatly between eligible patients (4); 2) LDCT screening can result in substantive harms and costs (1,5); 3) high competing (non-lung cancer) mortality and surgical treatment risk (1) can offset the all-cause mortality benefit from reducing lung cancer-specific mortality); and 4) LDCT screening is widely considered to be a preference-sensitive decision for at least some eligible individuals (6–8). A preference-sensitive decision occurs when the most appropriate decision for an individual depends on how they value the tradeoffs and risk (9). This complexity creates uncertainty about how to best promote high-benefit screening to patients, while respecting individual preferences.

Previous studies have highlighted the potential population benefits of risk-based screening (4,5,10–15). However, important practical questions remain. How can clinicians tailor recommendations to both clinical benefit and patient-preference? Given time-demands on primary care providers, should current demands for *detailed* shared-decision-making (required by CMS) be enforced even for patients who can expect a much larger (or lower?) than average mortality benefit? Clinicians need additional guidance to help them determine when screening is highly preference-sensitive versus when, if ever, it might be less preference sensitive. Understanding for whom LDCT screening is most preference-sensitive can help busy clinicians know when to prioritize their limited time for detailed shared

decision-making (16). To examine how to personalize recommendations for LDCT screening, we developed a microsimulation model that is consistent with large randomized trial data (1,17,18), can generate individual-specific estimates of the net benefit of screening, and allowed us to evaluate how sensitive the benefit is to different patient preferences.

Methods

Overview and Study Design

We created a markov microsimulation model (19) to examine clinical outcomes and health states that individual patients would experience under two screening scenarios, based on the NLST clinical trial: 6-year outcomes for 3 years of annual screening for lung cancer with low-dose computed tomography (LDCT) vs no LDCT screening (Figures 1s and 2s). The primary outcome was lifetime quality adjusted life-years (QALYs) from events occurring during the 6-year period. To examine the effect of preferences on the net benefit of screening, we varied disutilities across a likely range. Disutilities quantify the “degree of dislike.” In this study, disutilities quantify negative feelings about LDCT screening, screening outcomes, and follow-up care. These disutilities result in short term decrements in quality of life (Table 1). They are subtracted from a person’s baseline utility in each year that an event occurs (e.g., an LDCT screen or invasive procedure in that year). We derived “base-case” disutilities from a previously published decision analysis (20). However, our primary goal was to examine net benefit over a likely range of preferences (Table 1). Thus, understanding how quality adjusted life-year gains vary across the *range* of disutilities is much more important than understanding isolated outcomes under the uncertain point estimates used in the base-case. We also examined the effect of varying the rate of false-positives and overdiagnosis associated with LDCT screening.

US-Representative LDCT-Eligible Study Population

To estimate contemporary US-representative lung cancer screening outcomes, we used the National Health Interview Survey (NHIS) to simulate a nationally representative study sample of 1 million US Preventive Services Task Force-eligible heavy smokers (≥ 30 pack-years; current or former smoker quitting < 15 years ago) aged 55–80. The NHIS is an annual, cross-sectional representative, self-report survey of the noninstitutionalized US population (21). We first used harmonized NHIS data from the Integrated Health Interview Series (IHIS) to create an initial dataset of 14,422 respondents meeting current US Preventive Services Task Force eligibility criteria between 2010–2014. We then expanded the population based on NHIS sampling weights using R version 3.3.1 (22), performing multiple imputation for (and reporting) missing data (see Methods Supplement S.2).

The Microsimulation Model and Individualized Transition Probabilities

Non-screened cohort—Each individual smoker in the study population faced yearly transition probabilities of developing lung cancer (or not), dying from causes other than lung cancer (competing mortality), or dying from lung cancer (or not). We derived an individual’s competing mortality using a previously published model, which uses age, sex, pack-years, and smoking status (23,24). Estimates of individual-specific annual lung cancer incidence were based on the Bach et. al. validated model, which uses age, sex, smoking duration,

average cigarettes per day, and number of years since quitting (25) and is one of the best performing currently available lung cancer risk models (26) (all individuals were assumed to not have any asbestos exposure; see Methods Supplement S.3). We assigned histology (27) of incident lung cancers using a prediction model we developed from the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial control arm (18) (see S.3) and assigned lung cancer stage (stages Ia/b, II, IIIa/b, or IV) (28) using a stage distribution obtained from 2005–2012 Surveillance, Epidemiology, and End Results data, conditioned on histology and sex. To determine lung cancer survival, we used a Cancer Survival Analysis Software (29) cure survival model with lognormal distribution (30) fitted to Surveillance, Epidemiology, and End Results lung cancer data by sex, age group, histology and stage and then further calibrated to match NLST lung cancer mortality. Table 1 describes model transition probabilities and provides references. Additional model details are provided in the Methods Supplement S.3.

Screened cohort—The screened cohort experienced 3 rounds of annual LDCT screening (LDCT). Persons screened experienced a positive or a negative LDCT scan. All positive results underwent follow-up testing and could experience invasive diagnostic procedures, complications, and diagnostic mortality at the same rates reported in the NLST (1,17). A positive screen could represent a false positive finding (the person was not found to have lung cancer after further diagnostic testing) or a true positive that detected lung cancer. Each of the following could occur for persons with lung cancer detected at an earlier preclinical stage: stage-shift (detection and treatment prolonged time to lung cancer death); cure (prevention of lung cancer death due to curative lung resection) (35); or overdiagnosis (diagnosis of asymptomatic cancer that would never cause clinical disease or mortality in the absence of screening) (36). Other persons experienced a negative screen, which could represent a true negative (no lung cancer present) or a false-negative screen (lung cancer was present but missed due to the imperfect sensitivity of LDCT screening).

We based the sensitivity of the LDCT screen on the sensitivity reported in previously published work (31), calibrating to the lung cancer incidence and mortality observed in the NLST. We based the specificity of each round of LDCT screening on NLST results (1). To account for overdiagnosed lung cancers in our model, we retrospectively increased the number of lung cancers diagnosed in the screened cohort such that overdiagnosed cases represented 10% of all screen-detected lung cancers -- based on prior estimates from NLST and prior modeling studies (24,37,38) (see Methods Supplement S.3).

Model Assumptions—We assumed a 95% adherence rate for LDCT screening for the risk stratified analysis. We assigned overdiagnosis cases preferentially to individuals with a higher risk for developing clinical lung cancer, since most are probably true but very slowly growing cancer cases. We also conducted a sensitivity analysis that assigned overdiagnosis cases randomly across all patients.

Quality Adjusted Life Expectancy

We obtained utility (5,33) and disutility (“degree of dislike”) (20) estimates from previously published models (Table 1). We also assigned a disutility or degree-of-dislike to each

screening and follow-up imaging test, (a quality of life deduction of 0.4 days or 9.6 hours per LDCT or follow-up imaging in the base-case). This deduction in quality-adjusted life years accounted for the negative consequences of having the screening/testing itself (including the time and travel costs, out-of-pocket cost, and non-zero patient burden of undergoing LDCT screening/follow-up imaging) -- and also the potential negative psychological effects (e.g., distress from an indeterminate nodule) (see Methods Supplement S.3). Because of the inherent uncertainty in patient preferences, our primary aim in this study was to evaluate how varying personalized disutility (degree-of-dislike) estimates over a broad range influences the net benefit of LDCT screening (Table 1).

Outcomes and Uncertainty Analysis

We first examined the effect of a person's annual lung cancer risk on 1) number needed to screen with LDCT to avoid 1 lung cancer death and 2) QALY gains from LDCT versus no screening. Next, we examined risk-based outcomes for "pro-screening" patients with favorable preferences (i.e., lower screening disutilities or minimal negative feelings). We then examined risk-based outcomes for "anti-screening" patients with unfavorable preferences (i.e., higher screening-related disutilities or stronger negative feelings). To provide clinical context, we performed these analyses for each percentile of baseline lung cancer risk ("risk-based stratification") and for 4 specific individual patient scenarios ("individual attribute" analysis); Table 2). We also examined QALY gains for each quintile of competing mortality risk. A 3% discount rate was applied to all QALY calculations. We estimated the uncertainty in our results by calculating 95% bootstrap uncertainty ranges and performing sensitivity analyses on key parameters.

Sensitivity Analysis

Because NLST examined LDCT vs. chest x-ray (CXR) and found a 20% relative risk reduction (RRR), the RRR of LDCT vs. *no screening* is uncertain. In our base-case model, the RRR with LDCT screening vs. no screening is an output that is based largely on the sensitivity assigned to LDCT (31). We hypothesized that the RRR of LDCT vs. no screening in our base-case would be slightly larger than the 20% RRR with LDCT vs. CXR, reflecting a small benefit from annual CXR (e.g., due to true and false positive CXRs leading to CT). Nonetheless, we also planned for and conducted sensitivity analysis calibrating to the NLST's 20% RRR, which would reflect the RRR of LDCT screening vs. no screening if there is no benefit from annual CXR. Given substantial uncertainty about real-world rates of false-positive LDCT (suspicious results eventually found to not be cancer) (3,39) -- and also uncertainty about the true rate of overdiagnosis (diagnoses of "cancers" that never would have harmed the patient) (18,24,37,38) -- we tested the impact of varying these rates over a broad range (10% to 60% rate of false-positives for each LDCT screen; rate of overdiagnosis ranged from 5% to 15% of all lung cancers found).

This study was conducted without the support of a primary funding source.

Results

Lung Cancer Diagnoses and Deaths

Demographic attributes of the US-representative study population are detailed in Table 11s. Our model estimates aligned well with lung cancer incidence, mortality, and stage-distributions observed in the NLST and Prostate, Lung, Colorectal, and Ovarian Cancer Screening trials (see Figures 6s–7s and Tables 9s and 10s). When compared with no screening, our simulations showed a 22.7% RRR in lung cancer mortality with 3 LDCT screens vs. no screening, after 6.5 years of follow-up (95% uncertainty range 22.2%–23.3%). Given the 20% RRR in lung cancer mortality observed in the NLST (95% uncertainty range 6.8% to 26.7%), our finding suggests that CXR screening may have a small benefit.

Heterogeneity in QALY Gains by Baseline Risk (Risk-based Stratification)

Three rounds of annual LDCT screening led to 2,700 lifetime QALYs gained per 100,000 persons or 10 days of quality adjusted life per person (95% uncertainty range 2,600–2,800 per 100,000) among the US-representative population. However, QALY gains varied substantially as a function of a person's risk for developing lung cancer (Figure 1). QALY gains ranged from 830 per 100,000 in lower lung cancer risk persons to 3,500 per 100,000 in higher risk persons in our base-case. The number needed to screen (NNS) to avoid 1 all-cause death also varied substantially across risk groups, from an NNS of 537–572 for the lowest risk decile of eligible persons to an NNS of 95–98 in the highest risk decile. Increasing QALY gains peaked once lung cancer risk neared the 70th percentile and thereafter began to decline. This reduced net benefit in the highest lung cancer risk group is mainly due to reduced life-expectancy from competing risks (Results Supplement S.7). Although patient preferences also had a major impact on the net benefit of LDCT screening (Figure 1), the average person with lung cancer risk between 28th and 91st percentile for the screen eligible population (i.e., annual lung cancer risk between 0.3% and 1.3%), experienced net benefit even assuming preferences highly unfavorable toward screening.

Competing mortality risk

As expected, when competing mortality risk (dying from non-cancer causes) increased, the benefit (incremental QALY gains) from LDCT screening declined (Figure 2, Panel A). The decline was most striking for those in the highest competing risk quintile (life-expectancy < 10.5 years), where LDCT screening was preference-sensitive across all percentiles of lung cancer risk (Figure 2, Panel B). On the other hand, for those in the 4 lower quintiles of competing risk (life-expectancy > 10.5 years), the net benefit of LDCT screening continued to increase with lung cancer risk and was never preference-sensitive for annual lung cancer risk > 0.3% (Figure 2, Panel C).

Risk Thresholds: Preference-Sensitive and Preference-Insensitive for LDCT Screening

Putting together the results above, we find that the clinical benefit of LDCT depends on a person's life expectancy and annual risk of lung cancer. For the ~52.9% in our US representative population of patients with sufficiently long life-expectancy (median 10.5

years or greater) and high annual lung cancer risk (0.3% or higher per year), preferences against screening in the range considered in the analysis would diminish the benefit of screening but would not alter its superiority (net positive QALY benefit). For a few eligible patients (<0.1%) with both limited life expectancy (<10.5 years) and low (<0.3%) annual lung cancer risk, our base-case demonstrated possible net harm. For roughly 47%, the optimal strategy depended on individual patient preferences over the ranges considered in the models.

Individual Scenarios

To provide additional clinical context, we simulated expected outcomes for 4 persons, each with a specific set of risk attributes, and present how differences in preferences would play a role in determining quality-adjusted life-day gains or losses for each person (Table 3).

Sensitivity Analyses

Calibrating to a 20% RRR in mortality with LDCT screening vs. no screening, we found screening was preference-sensitive up to a 0.5% annual lung cancer risk, rather than the 0.3% risk threshold observed in our base-case (Table 12s). Varying the false-positive rate from 10%–60% was not highly influential in determining either the net benefit or preference-sensitivity of screening (see supplement). Varying the overdiagnosis rate from 5%–15% did not have a large impact on net benefit for most individuals, so screening remained preference-sensitive up to ~0.3% annual lung cancer risk, similar to our base-case. However, the rate of overdiagnosis was potentially influential in determining net benefit and preference-sensitivity among those at highest risk of developing lung cancer (due to the high competing mortality risk and shorter life-expectancy in this very high-risk group) (Results Supplement S.8). Table 12s summarizes how all sensitivity analyses impacted the preference-sensitive risk thresholds.

Discussion

We found that the health gains of LDCT lung cancer screening vary considerably across the eligible population, with three factors being highly influential: risk of lung cancer, competing risks/life expectancy, and patient preferences. For eligible persons with annual lung cancer risk > 0.3% and life expectancy > 10 years, LDCT screening overcame even highly negative views about screening and its downsides. For this high benefit group, which represented roughly 50% of our US-representative screen-eligible study population, the expected absolute mortality benefit was greater than that for most other routinely recommended cancer screening interventions (40,41). For eligible persons at lower lung cancer risk or more limited life expectancy, LDCT screening was usually highly preference sensitive. These results suggest an approach that may help clinicians improve and personalize LDCT screening discussions, after first assessing the net benefit of screening for their individual patients. This personalized assessment of net benefit should focus on using a validated lung cancer risk prediction tool to calculate a person's lung cancer risk (42,43) and also carefully consider the person's life-expectancy (44).

Interestingly, we also found that widely varying rates of false-positive findings (between 10%–60%) made only small differences in our results. Even when using large patient disutilities (i.e., stronger negative feelings) for false positive findings, we found little impact on net benefit for higher risk screening-eligible persons. This finding is of particular importance given a recent report that false positive LDCTs may be much higher in routine care than in the clinical trials (3). The insensitivity to small screening- or treatment-related disutilities when the absolute risk reduction (ARR) is high (i.e., absolute benefit is high) has been a consistent finding in other modeling (45,46). On the other hand, we found that a rate of overdiagnosed lung cancers can be an influential factor in the screening decision for a small group of eligible persons at very high risk of developing lung cancer, but only if overdiagnosis is strongly correlated with clinical lung cancer risk.

Previous research has demonstrated that LDCT screening was much more effective for those in higher quintiles of lung cancer risk (4,5,10–15). A recent study also suggests that current smoking trends could enhance the advantages of risk-based LDCT screening over time (47). Our objective was to move beyond the goal of optimizing risk-based eligibility criteria toward substantively informing a question at the core of delivering patient-centered care: “How should I present the pros and cons of screening for *this* eligible patient?” Our findings are thus highly relevant for those interested in taking a nuanced, patient-centered approach to discussing lung cancer screening. First, our results suggest that for some high-risk, high-benefit persons such as Person C in Table 2, lung cancer screening could generally be recommended during shared decision making (i.e., persons with annual lung cancer risk >0.3% and good life expectancy). For such high-benefit persons, LDCT screening’s absolute mortality benefit is on a par with that of average-risk colon cancer screening (for example, NNS to avoid a cancer-specific death of ~130 for *3 annual LDCT screens* among persons in the 7th decile of lung cancer risk vs. NNS of ~50 for sigmoidoscopy *every 5 years over a 25 year period*) (40); and LDCT is considerably more effective than screening mammography (NNS of ~1,000 to avoid a breast cancer death with *10 years of repeat screening mammography* among women ages 50–59) (41). An important caveat is that the large absolute mortality benefit of LDCT screening among high-risk persons will only translate into substantial QALY gains when the person also has good life-expectancy.

Of course, the initial screening recommendation can be overturned by the patient. Patients always have veto power and the authority to make the final screening decision. This autonomy means that high benefit patients with very strong negative preferences have the right to decide against LDCT screening, even if initially recommended. Such decisions should be respected, just as when patients refuse colon cancer screening or other recommended interventions.

Second, our results demonstrate that LDCT screening is highly preference-sensitive for some eligible persons with a lower absolute mortality benefit (annual lung cancer risk < 0.3% or life-expectancy < 10 years). Shared decision-making is most important for this group but is made complex by the harm/benefit tradeoffs being highly influenced by several factors. This complexity makes it all but impossible for clinicians to estimate an individual patient’s harm/benefit trade-offs without a decision-tool. Multiple online tools are available (42,43,48–50), including patient-facing tools (43) and a provider-facing tool we developed

using these study findings (42). To highlight how our results can help those counseling eligible persons about lung cancer screening, we summarize our findings in the form of simple, practical rules of thumb for clinicians in Table 3.

Limitations

First, currently available risk prediction models use different individual attributes to estimate lung cancer risk (e.g., race, family history, body mass index), which can lead to substantially different absolute risk estimates for the same person (26). Also, for a given population, agreement between predicted and observed lung cancer risk (calibration) will vary across risk models. This variation can lead to misclassification across a decision threshold, such as determining when a person is in a high-benefit vs. preference-sensitive category. Still, other methods of categorizing individual patients will likely lead to greater misclassification. Our approach, using a reasonably accurate risk prediction model to determine a person's degree of benefit, likely minimizes misclassification errors compared to other methods (e.g., using current USPSTF eligibility criteria for LDCT screening can lead to misclassification of high-risk persons who are 54 years old) (10–14). Also, any method that uses detailed smoking history will lead to misclassification due to unavoidable measurement error (e.g., caused by under- or over-reporting smoking history information or patient misunderstanding about what average packs-per-day means). Second, the risk thresholds we identified in this study, which are based on a person's 'true' lung cancer risk, are likely to vary depending on the structure of the microsimulation model used. Thus, the rules of thumb presented in Table 3 should be taken as rough guide-posts and not as rigid cutoffs.

Third, we modeled a 95% rate of adherence for all individuals, which is higher than is likely to occur. While considering adherence is appropriate for a policy analysis of population effects, examining net benefit assuming good adherence is a reasonable way to examine our question of how to personalize screening discussions. Fourth, our findings are specific to 3 annual screens, as we wanted to use parameter estimates from the NLST, which studied only 3 screening exams. The cumulative benefit of screening would be different with longer periods of annual screening but would require additional assumptions. In addition, limiting our analyses to outcomes from 3 annual screens is clinically useful because it is reasonable for clinicians to re-discuss the value of ongoing screening at least every 3–5 years with eligible patients. Not only does this allow for discussion of important new evidence on screening, it also allows clinicians and patients to discuss any changes to the patient's circumstances, underlying health, or preferences. Also, while we focused on the USPSTF eligible population, our results should extend beyond since both lung cancer and competing risk are taken into account. Fifth, we assumed the same rates of false-positive findings, invasive procedures, complications, etc. across the study population. This assumption is supported by prior studies demonstrating that rates of harm do not vary nearly as dramatically as the absolute mortality reduction (11). Moreover, substantial variation in the rates of false-positive findings did not substantially impact our results. Finally, we assumed all overdiagnosed lung cancers were stage IA adenocarcinomas. This assumption could underestimate the negative impact of a specific rate of overdiagnosis. However, we examined net benefit with screening over a broad range of overdiagnosis rates and also

assessed the impact of 2 different methods of assigning overdiagnosis; incorporating in this way uncertainty about overdiagnosis in the real world.

Conclusion

Our results support the importance of personalizing the harm/benefit assessment of LDCT lung cancer screening for informing screening decisions, rather than uniformly recommending or withholding a recommendation for eligible patients. Because the harm/benefit considerations can be complex, we have created and made available a web-based decision tool that incorporates the “rules of thumb” derived from our findings to facilitate personalized discussions about LDCT screening (42).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

none

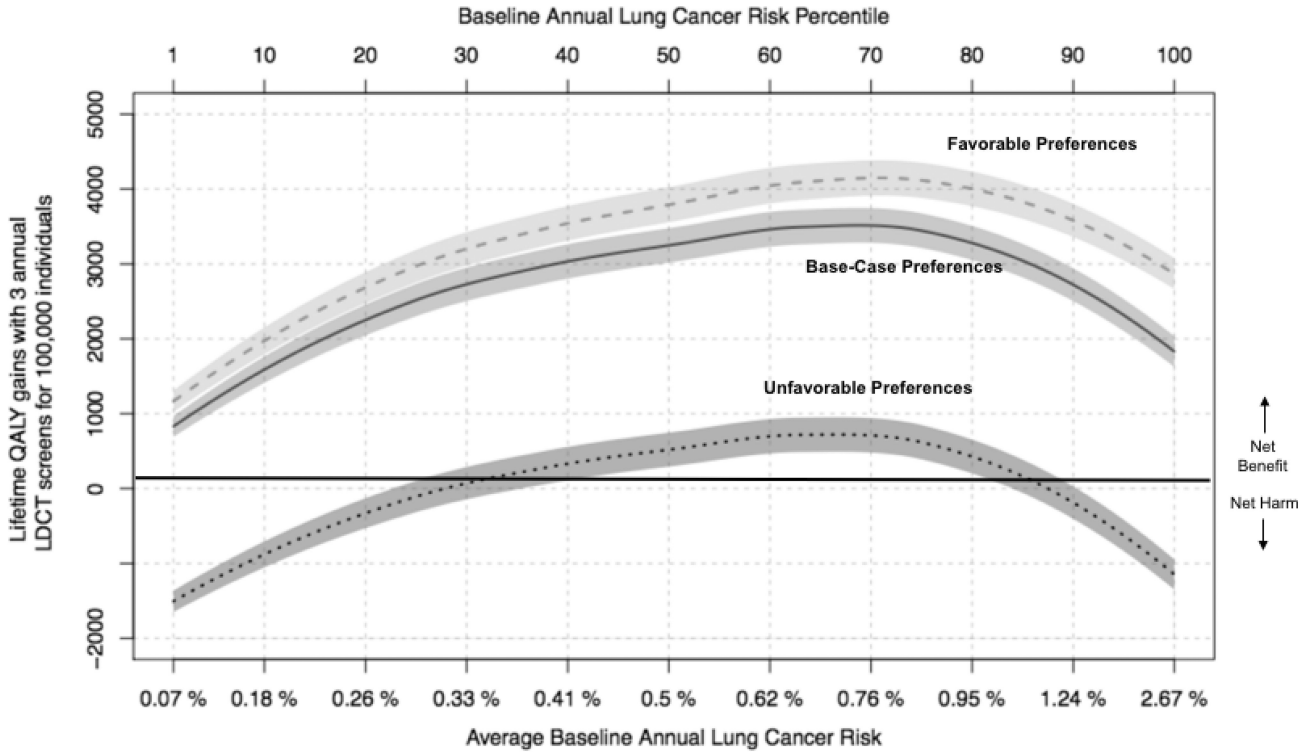
Grant Support: Development of the provider-facing web-tool using the results of this study was supported by a grant from the Veterans Affairs Quality Enhancement Research Initiative, QUE 15-286. Dr. Meza and Ms. Cao acknowledge support from NCI grant U01CA199284.

References

1. National Lung Screening Trial Research Team. Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening. *N Engl J Med*. 2011 Aug 4; 365(5):395–409. [PubMed: 21714641]
2. [cited 2018 Mar 11] Cancer of the Lung and Bronchus - SEER Stat Fact Sheets [Internet]. Available from: <http://seer.cancer.gov/statfacts/html/lungb.html>
3. Kinsinger LS, Anderson C, Kim J, Larson M, Chan SH, King HA, et al. Implementation of Lung Cancer Screening in the Veterans Health Administration. *JAMA Intern Med*. 2017 Mar 1; 177(3): 399–406. [PubMed: 28135352]
4. Kovalchik SA, Tammemägi M, Berg CD, Caporaso NE, Riley TL, Korch M, et al. Targeting of Low-Dose CT Screening According to the Risk of Lung-Cancer Death. *N Engl J Med*. 2013; 369(3):245–54. [PubMed: 23863051]
5. Black WC, Gareen IF, Soneji SS, Sicks JD, Keeler EB, Aberle DR, et al. Cost-Effectiveness of CT Screening in the National Lung Screening Trial. *N Engl J Med*. 2014 Nov 5; 371(19):1793–802. [PubMed: 25372087]
6. Moyer VA. Screening for Lung Cancer: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med*. 2014 Mar 4; 160(5):330–8. [PubMed: 24378917]
7. Sox HC. Implementing lung cancer screening under Medicare: The last chance to get it right? *JAMA*. 2014 Sep 24; 312(12):1206–7. [PubMed: 25247515]
8. Volk RJ, Hawk E, Bevers TB. Should CMS cover lung cancer screening for the fully informed patient? *JAMA*. 2014 Sep 24; 312(12):1193–4. [PubMed: 25247511]
9. Medicare C for, Baltimore MS 7500 SB, Usa M. [cited 2018 Mar 11] Beneficiary Engagement and Incentives Models: Shared Decision Making Model [Internet]. 2016. Available from: <https://www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2016-Fact-sheets-items/2016-12-08-2.html>
10. Tammemägi MC, Katki HA, Hocking WG, Church TR, Caporaso N, Kvale PA, et al. Selection Criteria for Lung-Cancer Screening. *N Engl J Med*. 2013; 368(8):728–36. [PubMed: 23425165]

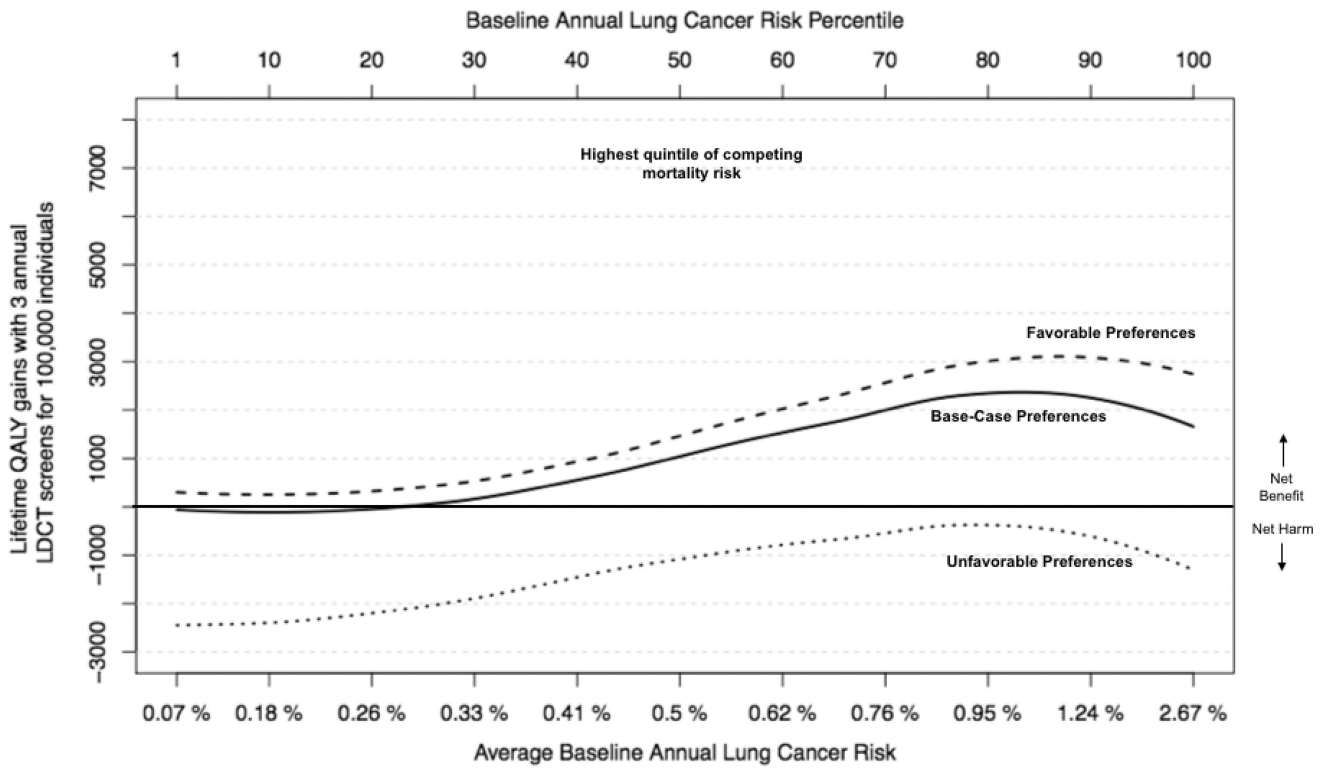
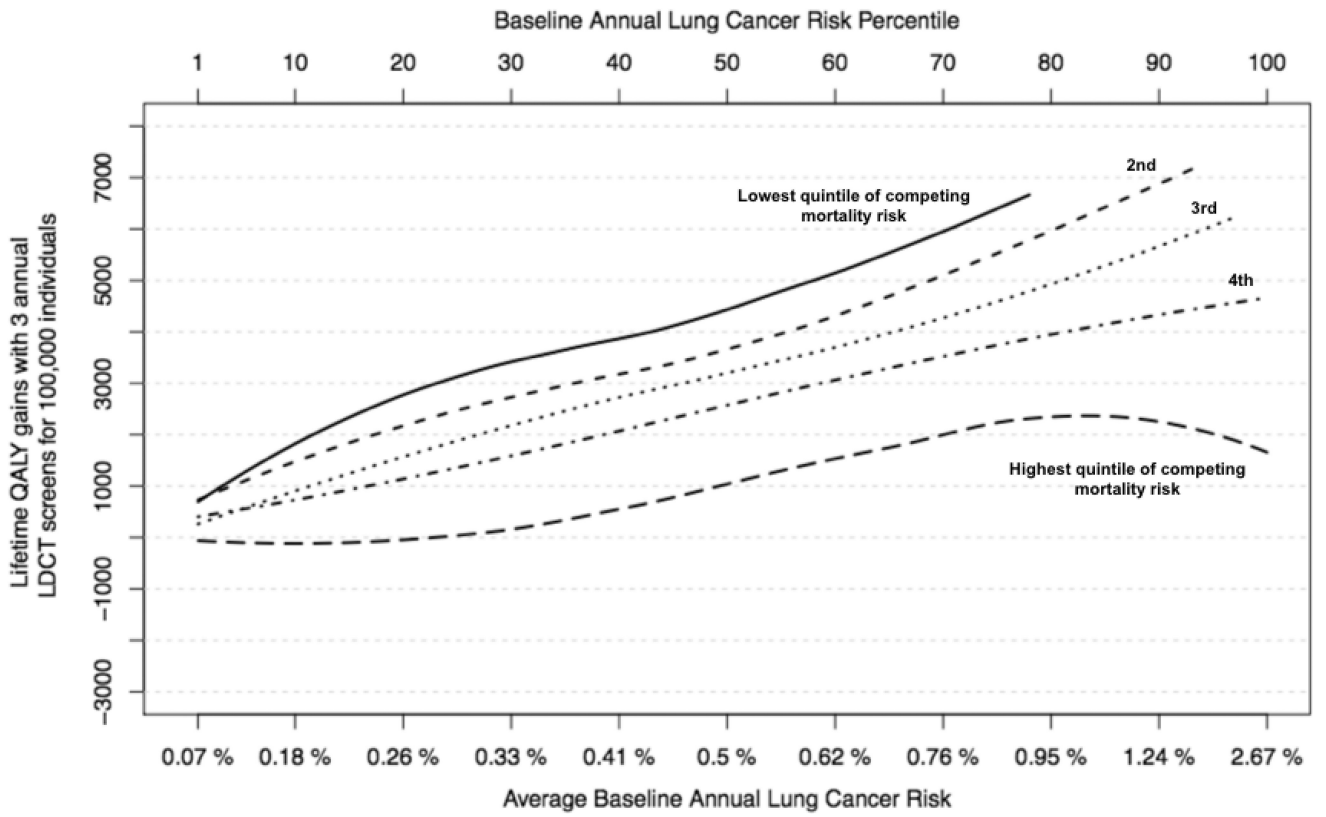
11. Katki HA, Kovalchik SA, Berg CD, Cheung LC, Chaturvedi AK. Development and validation of risk models to select ever-smokers for ct lung cancer screening. *JAMA*. 2016 Jun 7; 315(21):2300–11. [PubMed: 27179989]
12. Tammemägi MC, Church TR, Hocking WG, Silvestri GA, Kvale PA, Riley TL, et al. Evaluation of the Lung Cancer Risks at Which to Screen Ever- and Never-Smokers: Screening Rules Applied to the PLCO and NLST Cohorts. *PLoS Med*. 2014 Dec 2.11(12):e1001764. [PubMed: 25460915]
13. Ten Haaf K, Tammemägi MC, Bondy SJ, van der Aalst CM, Gu S, McGregor SE, et al. Performance and Cost-Effectiveness of Computed Tomography Lung Cancer Screening Scenarios in a Population-Based Setting: A Microsimulation Modeling Analysis in Ontario, Canada. *PLoS Med*. 2017 Feb.14(2):e1002225. [PubMed: 28170394]
14. Kumar V, Cohen JT, van Klaveren D, Soeteman DI, Wong JB, Neumann PJ, et al. Risk-Targeted Lung Cancer Screening: A Cost-Effectiveness Analysis. *Ann Intern Med* [Internet]. 2018 Jan 2.
15. Caverly TJ, Fagerlin A, Wiener RS, Slatore CG, Tanner NT, Yun S, et al. Comparison of Observed Harms and Expected Mortality Benefit for Persons in the Veterans Health Affairs Lung Cancer Screening Demonstration Project. *JAMA Intern Med* [Internet]. 2018 Jan 22.
16. Yarnall KS, Pollak KI, Østbye T, Krause KM, Michener JL. Primary care: is there enough time for prevention? *Am J Public Health*. 2003; 93(4):635–641. [PubMed: 12660210]
17. Aberle DR, DeMello S, Berg CD, Black WC, Brewer B, Church TR, et al. Results of the Two Incidence Screenings in the National Lung Screening Trial. *N Engl J Med*. 2013; 369(10):920–31. [PubMed: 24004119]
18. 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 Nov 2; 306(17):1865–73. [PubMed: 22031728]
19. Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-Transition Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. *Med Decis Making*. 2012 Sep 1; 32(5):690–700. [PubMed: 22990084]
20. Mahadevia PJ, Fleisher LA, Frick KD, Eng J, Goodman SN, Powe NR. Lung cancer screening with helical computed tomography in older adult smokers: A decision and cost-effectiveness analysis. *JAMA*. 2003 Jan 15; 289(3):313–22. [PubMed: 12525232]
21. [cited 2018 Mar 11] NHIS - National Health Interview Survey Homepage [Internet]. Available from: <https://www.cdc.gov/nchs/nhis/>
22. [cited 2018 Mar 11] R-3.3.1 for Windows (32/64 bit). [Internet]. Available from: <https://cran.r-project.org/bin/windows/base/old/3.3.1/>
23. Meza R, ten Haaf K, Kong CY, Erdogan A, Black WC, Tammemagi MC, et al. Comparative analysis of 5 lung cancer natural history and screening models that reproduce outcomes of the NLST and PLCO trials. *Cancer*. 2014 Jun 1; 120(11):1713–24. [PubMed: 24577803]
24. de Koning HJ, Meza R, Plevritis SK, ten Haaf K, Munshi VN, Jeon J, et al. Benefits and Harms of Computed Tomography Lung Cancer Screening Strategies: A Comparative Modeling Study for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2014 Mar 4; 160(5):311–20. [PubMed: 24379002]
25. Bach PB, Kattan MW, Thornquist MD, Kris MG, Tate RC, Barnett MJ, et al. Variations in lung cancer risk among smokers. *J Natl Cancer Inst*. 2003; 95(6):470–478. [PubMed: 12644540]
26. Haaf, K ten, Jeon, J., Tammemägi, MC., Han, SS., Kong, CY., Plevritis, SK., et al. Risk prediction models for selection of lung cancer screening candidates: A retrospective validation study. *PLOS Med*. 2017 Apr 4.14(4):e1002277. [PubMed: 28376113]
27. Fritz, AG., editor. International classification of diseases for oncology: ICD-O. Third. Geneva: World Health Organization; 2013. 242 p.First revision
28. [cited 2018 Mar 11] AJCC - Cancer Staging Manual [Internet]. Available from: <https://cancerstaging.org/references-tools/deskreferences/Pages/default.aspx>
29. [cited 2018 Mar 11] Cansurv - Surveillance Research Program [Internet]. Available from: <https://surveillance.cancer.gov/cansurv/>
30. Othus M, Barlogie B, LeBlanc ML, Crowley JJ. Cure Models as a Useful Statistical Tool for Analyzing Survival. *Am Assoc Cancer Res*. 2012 Jul 15; 18(14):3731–6.

31. Ten Haaf K, Rosmalen J van, Koning HJ de. Lung cancer detectability by test, histology, stage and gender: estimates from the NLST and the PLCO trials. *Cancer Epidemiol Biomarkers Prev*. 2014 Oct 13. cebp.0745.2014.
32. [cited 2017 Apr 12] SEER Data & Software for Researchers [Internet]. Available from: <https://seer.cancer.gov/resources/>
33. Tramontano AC, Schrag DL, Malin JK, Miller MC, Weeks JC, Swan JS, et al. Catalog and Comparison of Societal Preferences (Utilities) for Lung Cancer Health States Results from the Cancer Care Outcomes Research and Surveillance (CanCORS) Study. *Med Decis Making*. 2015 Apr 1; 35(3):371–87. [PubMed: 25670839]
34. Earle CC, Chapman RH, Baker CS, Bell CM, Stone PW, Sandberg EA, et al. Systematic overview of cost-utility assessments in oncology. *J Clin Oncol*. 2000; 18(18):3302–3317. [PubMed: 10986064]
35. Wever EM, Draisma G, Heijnsdijk EAM, de Koning HJ. How Does Early Detection by Screening Affect Disease Progression?: Modeling Estimated Benefits in Prostate Cancer Screening. *Med Decis Making*. 2011 Jul 1; 31(4):550–8. [PubMed: 21406620]
36. [cited 2018 Mar 11] NCI Dictionary of Cancer Terms [Internet]. Available from: <https://www.cancer.gov/publications/dictionaries/cancer-terms>
37. Patz EF Jr, Pinsky P, Gatsonis C, et al. Overdiagnosis in low-dose computed tomography screening for lung cancer. *JAMA Intern Med*. 2013 Dec 9.
38. Han SS, ten Haaf K, Hazelton WD, Munshi VN, Jeon J, Erdogan SA, et al. The impact of overdiagnosis on the selection of efficient lung cancer screening strategies: Overdiagnosis and lung cancer screening. *Int J Cancer*. 2017 Jun 1; 140(11):2436–43. [PubMed: 28073150]
39. Pinsky PF, Gierada DS, Black W, Munden R, Nath H, Aberle D, et al. Performance of Lung-RADS in the National Lung Screening Trial A Retrospective Assessment Performance of Lung-RADS in the NLST. *Ann Intern Med*. 2015 Apr 7; 162(7):485–91. [PubMed: 25664444]
40. Bibbins-Domingo K, Grossman DC, Curry SJ, Davidson KW, Epling JW, García FAR, et al. Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA*. 2016 Jun 21; 315(23):2564–75. [PubMed: 27304597]
41. Siu AL. on behalf of the U.S. Preventive Services Task Force. Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med*. 2016 Feb 16.164(4): 279. [PubMed: 26757170]
42. [cited 2018 Mar 11] Lung Decision Precision Provider Tool [Internet]. Available from: <https://share.lungdecisionprecision.com/>
43. [cited 2018 Mar 11] shouldiscreen web-tool [Internet]. Lung cancer screening decision aid. Available from: <http://www.shouldiscreen.com/>
44. [cited 2018 Mar 11] ePrognosis [Internet]. Available from: <http://eprognosis.ucsf.edu/>
45. Hayward RA, Kent DM, Vijan S, Hofer TP. Multivariable risk prediction can greatly enhance the statistical power of clinical trial subgroup analysis. *BMC Med Res Methodol*. 2006 Apr 13.6:18. [PubMed: 16613605]
46. Kent DM, Hayward RA. Limitations of applying summary results of clinical trials to individual patients: The need for risk stratification. *JAMA*. 2007 Sep 12; 298(10):1209–12. [PubMed: 17848656]
47. Cheung LC, Katki HA, Chaturvedi AK, Jemal A, Berg CD. Preventing Lung Cancer Mortality by Computed Tomography Screening: The Effect of Risk-Based Versus U.S. Preventive Services Task Force Eligibility Criteria, 2005–2015. *Ann Intern Med*. 2018 Feb 6.168(3):229. [PubMed: 29297008]
48. [cited 2018 Mar 11] Lung Cancer Screening Decision Tool | Memorial Sloan Kettering Cancer Center [Internet]. Available from: <http://nomograms.mskcc.org/Lung/Screening.aspx>
49. Lung Cancer Risk Calculators [Internet]. Brock University; Available from: <https://brocku.ca/lung-cancer-screening-and-risk-prediction/risk-calculators/> [cited 2018 Mar 11]
50. [cited 2018 Mar 11] Risk-based NLST Outcomes Tool (RNOT) | National Cancer Institute [Internet]. Available from: <https://analysistools.nci.nih.gov/lungCancerScreening/#/>



Mean Value	Lowest risk decile	2 nd	3 rd	4 th	5 th	6 th	7 th	8 th	9 th	Highest risk decile
NNS* (all-cause death)	537-572	325-340	255-267	211-220	183-190	159-165	138-143	123-127	112-115	95-98

Figure 1. Quality adjusted life year (QALY) gains with 3 annual low-dose computed tomography (LDCT) screens and lifetime follow-up (y-axis) by baseline lung cancer risk (x-axis). Baseline lung cancer risk is by percentile from 0–100. The table immediately below presents, for each corresponding decile of baseline risk, the number needed to screen to avoid 1 all-cause death.
 *NNS = Number needed to screen to avoid 1 death (95% bootstrap uncertainty range for mean value within each decile)



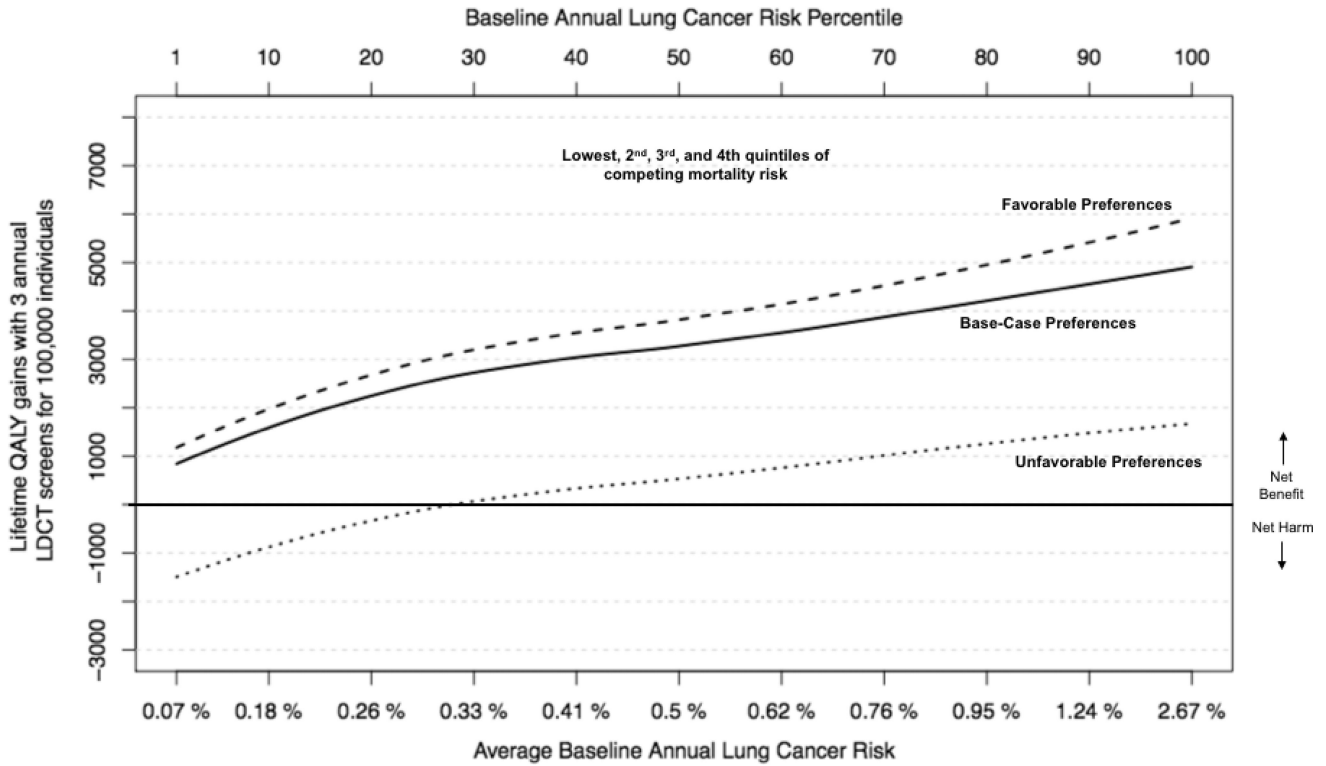


Figure 2.

Panel A. For each quintile of competing mortality risk, quality adjusted life year (QALY) gains (y-axis) by baseline lung cancer risk (x-axis) under base-case assumptions. Baseline lung cancer risk is by percentile from 0–100. Incremental QALY gains from LDCT screening decline as competing mortality risk increases.

Panel B. Quality adjusted life year (QALY) gains for highest (5th) quintile of competing mortality risk, by baseline lung cancer risk (x-axis).

Panel C. Quality adjusted life year (QALY) gains when excluding persons in highest (5th) quintile of competing mortality risk, by baseline lung cancer risk (x-axis).

Table 1

Model Parameters, Utilities, and Sources

Parameter	Description	Source
Lung Cancer Incidence	Bach et. al. annual lung cancer incidence prediction model	(25)
Lung Cancer Histology	Prediction model developed for this study and derived from control arm of the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial	(18) (see supplement for model development and accuracy)
Lung Cancer Stage	Using Surveillance, Epidemiology, and End Results data, conditioned on gender and histology	(23,24)
Lung Cancer Detectability	Sensitivity of LDCT screen by stage, histology, and screening round; calibrated to National Lung Screening Trial (NLST) CT arm	(31)
Lung Cancer-Specific Mortality	Using Surveillance, Epidemiology, and End Results data and Cancer Survival Analysis Software, calibrated to NLST, and conditioned on gender, age group, histology and stage	(1,29,32)
Other-Cause Mortality	By age, sex, pack-years, and smoking status using a published model	(23,24)
Baseline and Lung Cancer Utilities	From National Lung Screening Trial cost-effectiveness analysis and the Cancer Care Outcomes Study	(5,33)
Diagnostic tests, Diagnostic Complications, and Treatment	By stage, as observed in NLST	(1)

Disutilities (i.e., “degree of dislike” as deductions in days of quality-adjusted life): Base-case and range of values used in primary analyses (20)

Parameter	Base-Case (20)*	Favors Screening*	Against Screening*
Screening & Follow-up Imaging	-0.365 days	-0.0365 days	-3.65 days
Invasive Procedures	-10.95 days	-3.65 days	-18.25 days
Minor Complication	-3.65 days	-1.825 days	-7.3 days
Intermediate Complication	-10.95 days	-3.65 days	-21.9 days
Major Complication	-65.7 days	-18.25 days	-83.95 days
Surgery/Post-op period	-65.7 days	-18.25 days	-83.95 days
Radiation and/or Chemotherapy ^γ	-51.1 days	-69.35 days	-10.95 days

* Disutilities result in short term decrements in quality of life. They are subtracted from a person’s baseline utility in each year that an event occurs. With the exception of screening and follow-up imaging, our disutilities and their ranges were the same as a prior cost-utility analysis (20), which derived values for disutilities using a systematic review of cost-utility assessments in oncology (34). We assigned disutility to the screening and follow-up test as discussed in the text.

^γRadiation/chemotherapy works opposite from all other procedures/complications; a larger decrement in quality of life for radiation/chemotherapy (i.e., a more negative disutility) favors screening. This is because fewer patients in the screened cohort will be diagnosed with late-stage cancer due to the stage-shift with screening. Thus, fewer in the screened cohort will receive radiation/chemotherapy compared to no screening, so a higher chemo/radiation disutility advantages screening.

Scenarios Presenting Expected Outcomes[§] for 4 Individuals with Different Risk Profiles[‡] and Whether Net Benefit is Sensitive to Variation in Preferences

Table 2

Individual	Overall Risk Profile (annual lung cancer risk)	Median Life Expectancy (years)	Incremental Quality Adjusted Life-Days Gained (Base-case)	Incremental Quality Adjusted Life-Days (Favor-screening)	Incremental Quality Adjusted Life-Days (Anti-screening)	Number Needed to Screen*	False positives averted	Preference Sensitive?
Person A	60-year-old male at low risk (0.1%)	23.5	4 days	5 days	-6 days	513	405	Yes
Person B	75-year-old female at low risk (0.2%)	13.5	2 days	4 days	-8 days	355	279	Yes
Person C	60-year-old female at high risk (0.7%)	18.5	17 days	19 days	6 days	121	92	No
Person D	75-year-old male at high risk (1.2%)	7.6	5 days	8 days	-6 days	126	93	Yes

[§]Expected outcomes are for base-case assumptions regarding the sensitivity and specificity of low-dose computed tomography screening and assume a 10% rate of overdiagnosed lung cancers.

[‡]All individuals meet US Preventive Services Task Force eligibility criteria for low-dose computed tomography lung cancer screening.

* Number needed to screen to prevent 1 lung cancer death after 6.5-years of follow-up

Lower risk for developing lung cancer due to the following combination of characteristics: Former smoker; 30 pack-years (20 cigarettes per day for 30 years); quit 13 years ago; no asbestos exposure (in addition to these individual attributes, the following individual attributes were also used in assigning histology, which impacts incremental quality adjusted life-year gains with screening: body mass index of 25; No chronic obstructive pulmonary disease; No personal history of cancer; No family history of lung cancer).

Higher risk for developing lung cancer due to the following combination of characteristics: Current smoker; 67.5 pack-years (30 cigarettes per day for 45 years); no asbestos exposure (in addition to these individual attributes, the following individual attributes were also used in assigning histology, which impacts incremental quality adjusted life-year gains with screening: body mass index of 25; Positive history of chronic obstructive pulmonary disease; No personal history of cancer; Positive family history of lung cancer).

Table 3

Simple “rules of thumb” for personalizing LDCT screening discussions based on estimating a person’s annual lung cancer risk*

Screening is likely to be high benefit if a person’s...

- annual lung cancer risk is greater than ~0.3% and less than ~1.3%

Screening is likely to be highly preference-sensitive if a person’s...

- annual lung cancer risk is less than ~0.3% OR
- annual lung cancer risk is greater than ~1.3% (due to limited life-expectancy in this group) OR
- life-expectancy is limited (< 10.5 years)

Exercise caution if a person’s...

- annual lung cancer risk is very low (e.g., less than ~0.3%) AND their life-expectancy is limited (< 10.5 years). Screening may have negligible benefit or even net harm for these persons.
-

* All risk estimates are rounded to the nearest tenth of a decimal point. These values are presented as approximate “rules of thumb” to serve as guides to implementing individualized risk-based screening but should not be viewed as rigid cut-offs or “bright lines.” Clinicians can calculate a patient’s lung cancer risk using an online risk calculator that uses a well-validated lung cancer risk model (42). The Bach et. al. model (25) is one externally validated model that accurately calculates 1-year (annual) risk and was used in the provider-facing tool developed using the results of this study (42). The rules of thumb can be converted into estimates for 5-year, 6-year, and 10-year lung cancer risk thresholds to facilitate use of risk prediction models that use other time-frames:

a) 0.3% 1-year risk \cong 2.0% 5-year risk \cong 2.6% 6-year risk \cong 5.2% 10-year risk

b) 1.3% 1-year risk \cong 5.7% 5-year risk \cong 6.9% 6-year risk \cong 11.4% 10-year risk