



Published in final edited form as:

JAMA Intern Med. 2014 February 1; 174(2): 269–274. doi:10.1001/jamainternmed.2013.12738.

Overdiagnosis in Low-Dose Computed Tomography Screening for Lung Cancer

Edward F. Patz Jr, MD, Paul Pinsky, PhD, Constantine Gatsonis, PhD, JoRean D. Sicks, MS, Barnett S. Kramer, MD, MPH, Martin C. Tammemägi, PhD, Caroline Chiles, MD, William C. Black, MD, Denise R. Aberle, MD, and NLST Overdiagnosis Manuscript Writing Team

Department of Radiology, Duke University Medical Center, Durham, North Carolina (Patz); Department of Pharmacology and Cancer Biology, Duke University Medical Center, Durham, North Carolina (Patz); Division of Cancer Prevention, National Cancer Institute, Bethesda, Maryland (Pinsky, Kramer); Center for Statistical Sciences, Brown School of Public Health, Providence, Rhode Island (Gatsonis, Sicks); Department of Biostatistics, Brown School of Public Health, Providence, Rhode Island (Gatsonis); Department of Community Health Sciences, Brock University, St Catharines, Ontario, Canada (Tammemägi); Department of Radiology, Wake Forest University Health Sciences Center, Winston-Salem, North Carolina (Chiles); Department of Radiology, Dartmouth Medical School, Hanover, New Hampshire (Black); Department of Community and Family Medicine, Dartmouth Medical School, Hanover, New Hampshire (Black); Department of Radiology, University of California, Los Angeles (Aberle)

Copyright 2014 American Medical Association. All rights reserved

Corresponding Author: Edward F. Patz Jr, MD, Duke University Medical Center, Department of Radiology, Box 3808, Durham, NC 27710 (patz0002@mc.duke.edu).

Author Contributions: Dr Patz had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Patz, Gatsonis, Sicks, Kramer, Black, Aberle.

Acquisition of data: Gatsonis, Sicks, Tammemägi, Chiles.

Analysis and interpretation of data: Patz, Pinsky, Gatsonis, Sicks, Kramer, Tammemägi, Black.

Drafting of the manuscript: Patz, Pinsky, Gatsonis, Black.

Critical revision of the manuscript for important intellectual content: Patz, Pinsky, Gatsonis, Sicks, Kramer, Tammemägi, Chiles, Aberle.

Statistical analysis: Patz, Pinsky, Gatsonis, Sicks, Black.

Obtained funding: Gatsonis.

Administrative, technical, or material support: Gatsonis, Sicks.

Study supervision: Patz, Gatsonis, Kramer, Tammemägi.

Group Information: The members of the NLST Overdiagnosis Writing Team are Deni Aberle, MD, University of California, Los Angeles; Judith Amorosa, MD, Robert Wood Johnson University Hospital, East Brunswick, New Jersey; Christine Berg, MD, National Cancer Institute, Bethesda, Maryland; William C. Black, MD, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire; Caroline Chiles, MD, Wake Forest University Health Sciences Center, Winston Salem, North Carolina; Tim Church, PhD, MS, University of Minnesota, Minneapolis; David Crawford, MD, University of Colorado at Denver, Aurora; Richard Fagerstrom, PhD, National Cancer Institute, Bethesda, Maryland; Matthew T. Freedman, MD, Georgetown University, Washington, DC; Ilana F. Gareen, PhD, Brown University, Providence, Rhode Island; Kavita Garg, MD, University of Colorado at Denver, Aurora; Constantine Gatsonis, PhD, Brown University, Providence, Rhode Island; Barry Kramer, MD, MPH, National Cancer Institute, Bethesda, Maryland; David Lynch, MD, National Jewish Health, Denver, Colorado; Reginald F. Munden, MD, MD Anderson Cancer Center, Houston, Texas; Hrudaya (Bobby) Nath, MBBS, DMR, MD, University of Alabama, Birmingham; Edward F. Patz Jr, MD, Duke University Medical Center, Durham, North Carolina; Paul Pinsky, PhD, National Cancer Institute, Bethesda, Maryland; Mitchell Schnall, MD, PhD, University of Pennsylvania, Philadelphia; JoRean Sicks, MS, Brown University, Providence, Rhode Island; Martin C. Tammemägi, PhD, Brock University, St Catharines, Ontario, Canada; Joel Weissfeld, MD, MPH, University of Pittsburgh, Pittsburgh, Pennsylvania.

Correction: This article was corrected online March 10, 2014, for errors in the Conflict of Interest statement.

Abstract

IMPORTANCE—Screening for lung cancer has the potential to reduce mortality, but in addition to detecting aggressive tumors, screening will also detect indolent tumors that otherwise may not cause clinical symptoms. These overdiagnosis cases represent an important potential harm of screening because they incur additional cost, anxiety, and morbidity associated with cancer treatment.

OBJECTIVE—To estimate overdiagnosis in the National Lung Screening Trial (NLST).

DESIGN, SETTING, AND PARTICIPANTS—We used data from the NLST, a randomized trial comparing screening using low-dose computed tomography (LDCT) vs chest radiography (CXR) among 53 452 persons at high risk for lung cancer observed for 6.4 years, to estimate the excess number of lung cancers in the LDCT arm of the NLST compared with the CXR arm.

MAIN OUTCOMES AND MEASURES—We calculated 2 measures of overdiagnosis: the probability that a lung cancer detected by screening with LDCT is an overdiagnosis (P_S), defined as the excess lung cancers detected by LDCT divided by all lung cancers detected by screening in the LDCT arm; and the number of cases that were considered overdiagnosis relative to the number of persons needed to screen to prevent 1 death from lung cancer.

RESULTS—During follow-up, 1089 lung cancers were reported in the LDCT arm and 969 in the CXR arm of the NLST. The probability is 18.5% (95% CI, 5.4%–30.6%) that any lung cancer detected by screening with LDCT was an overdiagnosis, 22.5% (95% CI, 9.7%–34.3%) that a non-small cell lung cancer detected by LDCT was an overdiagnosis, and 78.9% (95% CI, 62.2%–93.5%) that a bronchioalveolar lung cancer detected by LDCT was an overdiagnosis. The number of cases of overdiagnosis found among the 320 participants who would need to be screened in the NLST to prevent 1 death from lung cancer was 1.38.

CONCLUSIONS AND RELEVANCE—More than 18% of all lung cancers detected by LDCT in the NLST seem to be indolent, and overdiagnosis should be considered when describing the risks of LDCT screening for lung cancer.

Screening for lung cancer has been proposed for decades. It is fundamentally based on the principle that tumors will be detected at a smaller size and earlier stage, when treatment is more effective, resulting in a reduction in lung cancer mortality.^{1,2} An ideal screening program targets individuals at the highest risk of lung cancer, uses a cost-effective test to detect tumors at an early stage, and efficiently excludes patients with clinically insignificant abnormalities.

Unfortunately, there is currently no ideal screening test for lung cancer, and a clear understanding of the risks and benefits should be considered in the design of a population-based screening program. Low-dose computed tomography (LDCT) has been suggested as a screening tool for lung cancer, and recent results from the National Lung Screening Trial (NLST) demonstrated an encouraging 20% relative reduction in lung cancer-specific mortality compared with screening using chest radiography (CXR).³ Whereas the decrease in mortality highlights the primary benefit of screening, the trial also found more cases of lung cancer in the LDCT group compared with the CXR group; this is a limitation of screening because some of these tumors may be indolent and clinically insignificant.

In previous CXR lung cancer screening trials, more lung cancers were detected in the screened arm than in the observational group.⁴⁻⁶ The excess number of early-stage lung cancers, even after extended follow-up, is usually attributed to overdiagnosis.^{7,8} Overdiagnosis is defined as the detection, usually by screening, of a cancer that would not otherwise have become clinically apparent; overdiagnosis is often an intrinsic feature of screening, which by definition seeks to detect occult disease in asymptomatic individuals. Overdiagnosis is 1 of the limitations of screening because it incurs unnecessary treatment, morbidity (and mortality in rare cases), follow-up, cost, and anxiety and labels a patient with a disease that otherwise would never have been detected.⁹

Estimating the true level of overdiagnosis in a screening trial such as the NLST requires a sufficiently long period of post-screening follow-up because there is typically a “catch-up” period in the nonscreened arm, during which more cases of cancer are diagnosed and potentially catch up to the greater number of cases detected earlier in the screened arm. The length of the potential catch-up period is related to the lead time associated with the screening modality, where the lead time is the difference between the time when diagnosis would have been made without screening and the time that the diagnosis was actually made as a result of early detection by screening.

In the NLST, there were 4 to 5 years of follow-up after screening, which may not be sufficient to detect all cancers in the control CXR arm; with longer follow-up, additional catch-up may occur. Therefore, the present study calculated rates of excess cancers in the LDCT vs CXR arm, for all lung cancer and for various histologic subtypes. These excess cancer rates provide an upper bound on the true overdiagnosis rate associated with LDCT screening relative to CXR screening.

To complement the descriptive analysis, we also developed a standard convolution-type model that when fit to the NLST data can be used to estimate overdiagnosis and excess cancer rates over various screening scenarios different from that used in the NLST.

Methods

Design and Overview of the NLST

A detailed description of the NLST design, methods, and initial results has been previously reported; the present study used extended follow-up data through December 31, 2009.³ From August 2002 through April 2004, 53 452 individuals at high risk for lung cancer, with at least a 30 pack-year history of cigarette smoking (former smokers had quit within the past 15 years), between the ages of 55 and 74 years were enrolled at 33 US medical centers into a prospective screening trial. The study protocol was approved by the institutional review board at each of the 33 screening centers, and written informed consent was obtained from each participant before randomization. All participants were randomly assigned to receive either 3 annual LDCT studies (26 722 participants) or 3 annual single-view posterior-anterior CXRs (26 730 participants) and then observed for up to an additional 5 years. The primary trial objective was to determine the effect of LDCT screening vs CXR screening on lung cancer mortality.

Lung cancer diagnoses were ascertained primarily through standardized forms administered to study participants at 6-month or 1-year intervals, which inquired about any recent cancer diagnoses. Trained abstractors confirmed reported cases of lung cancer using medical records and pathology reports. A screen-detected cancer was defined as a cancer diagnosed within 1 year of a positive screening result or diagnosed after a longer period on the basis of diagnostic procedures prompted by the screen.

Excess Cancer Cases

Here we use 2 definitions of an excess cancer rate, 1 emphasizing the clinical perspective (denoted P_S) and 1 emphasizing a more public health perspective (denoted P_A). For each, the numerator of the rate is the same, namely, the number of excess lung cancer cases in the LDCT arm as compared with the CXR arm, ie, the difference in the total count of lung cancers between the LDCT and CXR arms. For P_S , the denominator is the total number of screen-detected lung cancer cases in the LDCT arm. The quantity P_S is a measure of the probability that a participant's LDCT screen-detected cancer would not have become clinically apparent during the given screening phase if LDCT screening had not been performed. For P_A , the denominator is the total number of lung cancers diagnosed in the LDCT arm. Thus, P_A is the fraction of all lung cancer cases diagnosed during a given period in a cohort who underwent LDCT screening that would not have been diagnosed during that period absent the LDCT screening. We estimated these 2 indices using the observed lung cancer counts from the NLST; confidence intervals were obtained using bootstrapping. We also estimated the number of cases that were considered overdiagnosis, relative to the number of participants needed to screen to prevent 1 death from lung cancer.

Values of P_S and P_A were estimated for all lung cancer, all non-small cell lung cancer (NSCLC), bronchioloalveolar carcinoma (BAC), and NSCLC excluding BAC. As described in the Introduction, these excess cancer rate estimates represent an upper bound to true overdiagnosis rates because the post-screening follow-up period in the NLST may not have been long enough to totally differentiate overdiagnosis from the effects of lead time.

We estimated the number of overdiagnoses relative to the number needed to screen to prevent 1 lung cancer death as the number of excess lung cancer cases in the LDCT arm of the NLST divided by the difference in lung cancer deaths in the LDCT and CXR arms of the NLST.

Convolution Model Based on NLST Data

In addition to the aforementioned descriptive analysis of excess cancers from the NLST, we also fit a standard convolution model to the NLST data to estimate excess cancers relative to no screening and excess cancers expected if follow-up continued lifelong. The convolution model postulates a preclinical phase of disease and a mean sojourn time in that phase before clinical diagnosis; the model also estimates the sensitivity of screening (here with LDCT or CXR).^{10,11} See eAppendix (in Supplement) for more details.

Separate models were fit for BAC and non-BAC NSCLC. Simulations were then run using the fitted model parameters to compute excess cancer rates under various screening

regimens. Excess cancer rates P_A and P_S were defined similarly as in the descriptive analysis, although as stated, rates were computed both relative to CXR screening and relative to no screening (see eAppendix in Supplement for further details).

Results

Empirical Estimates of Overdiagnosis

The total numbers of lung cancer cases by year and by study arm are shown in Table 1. The mean follow-up in the LDCT arm was 6.41 years, and the mean follow-up in the CXR arm was 6.37 years. Table 2 shows the number of screen-detected and non-screen-detected cases of lung cancer in each arm according to general histological categories.

At the end of the entire trial, there were 1089 total lung cancer cases in the LDCT arm (649 detected by LDCT screening) and 969 cases in the CXR arm, for an excess of 120 cases. This gives excess cancer rates P_S and P_A of 18.5% (95% CI, 5.4%–30.6%) and 11.0% (95% CI, 3.2%–18.2%), respectively (Table 3). With respect to excess cancer rates for NSCLC, there were a total of 926 cases in the LDCT arm and 793 cases in the CXR arm, for a difference of 133 cases. Excess cancer rates were $P_S = 22.5%$ (95% CI, 9.7%–34.3%) and $P_A = 14.4%$ (95% CI, 6.1%–21.8%). There were 111 cases of BAC in the LDCT arm and 36 in the CXR arm, giving excess cancer rates of $P_S = 78.9%$ (95% CI, 62.2%–93.5%) and $P_A = 67.6%$ (95% CI, 53.5%–78.5%).

From the original NLST report,³ the number needed to screen to prevent 1 lung cancer death was 320. There were 443 and 356 lung cancer deaths in the CXR and LDCT arms, respectively, giving a difference of 87. As mentioned, there was an excess of 120 lung cancer cases in the LDCT compared with the CXR arm. Therefore, the number of cases of overdiagnosis found in the 320 participants needed to screen to prevent 1 death from lung cancer is $120/87 = 1.38$.

Model-Based Estimates

In general, the model fits the NLST data relatively well. The fit of the model, demonstrated by observed vs expected (modeled) case counts by mode of detection, time, and arm (LDCT vs CXR), is provided in eTable 1 (in Supplement).

The parameter estimates for NSCLC and BAC using the convolution model are shown in Table 4. The mean sojourn time for non-BAC NSCLC was 3.6 (95% CI, 3.0–4.3) years. This implies that 24.1% (42.4%) of cases would become clinically apparent within 1 year (2 years) of entering the preclinical phase (assuming no other-cause mortality). Estimated sensitivity was 83% (95% CI, 72%–94%) for LDCT and 33% (95% CI, 26%–42%) for CXR. For BAC, estimated sensitivity was 38% (95% CI, 7%–62%) for LDCT and 4% (95% CI, 1%–9%) for CXR. Mean sojourn time for BAC was 32.1 (95% CI, 17.3–270.7) years, with 14.4% (26.8%) of cases becoming clinically apparent in 5 (10) years.

Estimates of excess cancer rates (P_S) with LDCT under various screening scenarios are shown in Table 5. For the NLST scenario of 3 screens and roughly 7 years of total follow-up, the excess cancer rates for all NSCLCs were 31% vs no screening and 19% vs CXR. For

BAC under this scenario, rates were 85% (vs no screening) and 71% (vs CXR), whereas for non-BAC NSCLCs, rates were 21% (vs no screening) and 9% (vs CXR). For NLST screening (ie, 3 annual screens) but with lifetime follow-up, excess cancer rates decreased substantially, to 11% (vs no screening) and 9% (vs CXR) for all NSCLCs and 2.6% (vs no screening) and 1.2% (vs CXR) for non-BAC NSCLCs; rates for BAC were 49% and 41%, respectively. These latter rates, with lifetime follow-up, are thus estimates of actual overdiagnosis rates. Under a scenario of 5 annual screens and 5 years of total follow-up, excess cancer rates for non-BAC NSCLC were relatively high, 45% compared with no screening and 16% compared with CXR. However, with follow-up extended through participants' lifetimes, excess cancer rates were similar to those for 3 annual screens.

On the basis of the estimate for non-BAC NSCLC of a mean sojourn time (and mean lead time) of 3.6 years, approximately 25% of non-BAC NSCLC tumors would have lead times of longer than 5 years. This helps explain why the non-BAC NSCLC excess cancer rate decreased substantially when the follow-up was changed from 7 years total (5 years past the last screen) to lifetime follow-up.

Discussion

Screening for lung cancer with LDCT in the NLST showed a 20% relative reduction in mortality, and 320 participants were needed to screen to prevent 1 lung cancer death. These findings were met with enthusiasm, but before a widespread public health screening program is implemented, risks of screening also need to be considered. One of the limitations and potential harms is overdiagnosis because it is not clear that all early-stage lesions detected in asymptomatic individuals will progress to cause symptoms and affect long-term outcome. These patients may undergo an invasive diagnostic procedure, have surgical resection, be given a diagnosis of lung cancer, and require multiple sequential follow-up studies when some tumors are potentially clinically insignificant. These cases of overdiagnosis are treated as any other lung cancer because it is generally not possible to distinguish indolent lesions from more aggressive tumors.

The true extent of overdiagnosis in lung cancer is difficult to determine because most of what we know about this disease is derived from symptomatic patients.¹⁰ It is not possible to perform a trial that biopsies every pulmonary nodule and follows a randomized group of patients with lung cancer in an observational arm without therapy to determine the natural progression of the disease. However, there are studies that suggest some degree of overdiagnosis in lung cancer. First, prior screening trials with CXRs found an excess number of cancers in the screened arm, without a reduction in mortality.^{6,11,12} This excess of lung cancer cases was attributed to overdiagnosis. Second, autopsy studies have shown that patients die with undiagnosed lung cancer, and it is not the cause of death.^{13,14} Third, LDCT screening trials from Japan found that when the population of a geographic region was screened without discriminating on risk factors such as smoking, the lung cancer rates were often similar between smokers and non-smokers, suggesting that the more patients undergo imaging, the more tumors are found.^{15,16} And finally, a recent study used volume-doubling times on sequential LDCT to estimate overdiagnosis and suggested that approximately 25% of cases may be indolent.¹⁷

The present analyses were performed to provide an empirical estimate of, or at least an upper bound on, the magnitude of overdiagnosis in the NLST so that the impact on mass screening programs could be understood. As mentioned, the follow-up in the NLST may not have been long enough to account for the lead time of all LDCT-detected cancers, particularly because tumor growth rates are quite variable and do not consistently follow classical expected exponential growth curves.¹⁸ On the basis of the convolution model developed here, approximately 25% of non-BAC NSCLC tumors would have lead times of at least 5 years, or longer than the period from final NLST screen to end of follow-up. Thus, it is likely that some additional catch-up would occur in the NLST with longer follow-up because of CXR arm diagnosed cancers whose counterparts in the LDCT arm had a long lead time.

The data from this study suggest that, at most, 18% of persons in the LDCT arm with screen-detected lung cancer (P_5) and 22% of those in the LDCT arm with screen-detected NSCLC may be cases of overdiagnosis. In other words, if these individuals had not entered the NLST, they would not have received a lung cancer diagnosis or treatment, at least for the next 5 years. This is most striking in patients with a diagnosis of BAC. In the new International Association for the Study of Lung Cancer histologic classification of adenocarcinomas, many of these tumors would be designated as minimally invasive adenocarcinomas, suggesting an indolent behavior and good long-term outcome.¹⁹ These data raise the question as to the necessity and type of therapy required if a diagnosis of minimally invasive adenocarcinoma is established and challenge the diagnostic community to develop a classification scheme that could accurately phenotype all lung tumors.

In addition to the NLST data, the present study used a convolution model to explore the effect of other screening scenarios on overdiagnosis. The modeling generally provides a good fit to the NLST data and is useful for several reasons. First, it makes it possible to estimate excess cancers relative to no screening (and not just to CXR screening). Second, with the model, one can extrapolate to different screening and follow-up scenarios, not just what was observed in the NLST. Third, by extrapolating beyond the NLST follow-up period, the model can provide an estimate of actual overdiagnosis rates and not just an upper bound. Finally, the model may also be used to determine how overdiagnosis varies by lung cancer risk. Using the model, we found that the excess cancer rate associated with 3 annual LDCT screens decreased substantially in changing from the NLST follow-up (approximately 5 years after screening) to lifetime follow-up, with the latter estimates representing true overdiagnosis. There was a relatively low rate of true overdiagnosis for non-BAC NSCLC of approximately 3% (relative to no screening), although BAC still had high rates of overdiagnosis (approximately 50%).

The model also found that overdiagnosis rates with LDCT were greater relative to no screening than relative to screening with CXR. These data are consistent with prior screening trials using CXR and sputum cytologic analysis as compared with no screening in the standard of care group because the CXR group had an excess number of cases of lung cancer.

As with any model, one should be cautious in extrapolating much beyond the data on which the model was based, which in this case are 3 annual screens and a total of approximately 7 years of follow-up. Therefore, the estimates of true overdiagnosis, based on the lifetime follow-up scenarios, must be treated cautiously. Furthermore, the model does not always reflect true clinical practice because patients are not as predictable and reliable as the model system would suggest, particularly if different risk categories are explored. However, the model does provide a framework to explore a variety of screening parameters so that potential limitations of various screening conditions can be understood.

In summary, screening for lung cancer with LDCT has the potential to detect indolent tumors, resulting in overdiagnosis. Whereas the NLST demonstrated a relative mortality reduction with LDCT, the limitations of the screening process, including the magnitude of overdiagnosis, should be considered when guidelines for mass screening programs are constructed. In the future, once there are better biomarkers and imaging techniques to predict which individuals with a diagnosis of lung cancer will have more or less aggressive disease, treatment options can be optimized, and a mass screening program can become more valuable.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Conflict of Interest Disclosures: Dr Patz conducted a research project on serum protein biomarkers and indeterminate pulmonary nodules funded by Laboratory Corporation of America through a sponsored research agreement with Duke University. Dr Gatsonis is a consultant/ member of the scientific advisory board for Wilex AG and a board member for Frontier Science & Technology Research Foundation and has served as a consultant to Endocyte Inc and Genentech.

Funding/Support: This research was supported by the National Institutes of Health (grants U01 CA079778 and U01 CA080098 and contracts N01-CN-25511, N01-CN-25512, N01-CN-25513, N01-CN-25514, N01-CN-25515, N01-CN-25516, N01-CN-25518, N01-CN-25522, N01-CN-25524, N01-CN-75022, N01-CN-25476, and N02-CN-63300).

Role of the Sponsor: The National Institutes of Health had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

REFERENCES

1. Patz EF Jr, Goodman PC, Bepler G. Screening for lung cancer. *N Engl J Med.* 2000; 343(22):1627–1633. [PubMed: 11096172]
2. Bach PB, Mirkin JN, Oliver TK, et al. Benefits and harms of CT screening for lung cancer: a systematic review. *JAMA.* 2012; 307(22):2418–2429. [PubMed: 22610500]
3. National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med.* 2011; 365(5):395–409. [PubMed: 21714641]
4. Fontana T, Sanderson DR, Woolner LB, et al. The Mayo Lung Project for early detection and localization of bronchogenic carcinoma: a status report. *Chest.* 1975; 67:511–522. [PubMed: 1126186]
5. Fontana RS, Sanderson DR, Taylor WF, et al. Early lung cancer detection: results of the initial (prevalence) radiologic and cytologic screening in the Mayo Clinic study. *Am Rev Respir Dis.* 1984; 130(4):561–565. [PubMed: 6091507]

6. 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(16):1308–1316. [PubMed: 10944552]
7. Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst.* 2010; 102(9):605–613. [PubMed: 20413742]
8. Patz EF Jr. Lung cancer screening, overdiagnosis bias, and reevaluation of the Mayo Lung Project. *J Natl Cancer Inst.* 2006; 98(11):724–725. [PubMed: 16757691]
9. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. *N Engl J Med.* 2012; 367(21):1998–2005. [PubMed: 23171096]
10. Detterbeck FC. Cancer, concepts, cohorts and complexity: avoiding oversimplification of overdiagnosis. *Thorax.* 2012; 67(9):842–845. [PubMed: 22628119]
11. Fontana RS, Sanderson DR, Woolner LB, et al. Screening for lung cancer: a critique of the Mayo Lung Project. *Cancer.* 1991; 67((4)(suppl)):1155–1164. [PubMed: 1991274]
12. Marcus PM, Bergstralh EJ, Zweig MH, Harris A, Offord KP, Fontana RS. Extended lung cancer incidence follow-up in the Mayo Lung Project and overdiagnosis. *J Natl Cancer Inst.* 2006; 98(11):748–756. [PubMed: 16757699]
13. Dammas S, Patz EF Jr, Goodman PC. Identification of small lung nodules at autopsy: implications for lung cancer screening and overdiagnosis bias. *Lung Cancer.* 2001; 33(1):11–16. [PubMed: 11429191]
14. McFarlane MJ, Feinstein AR, Wells CK. Necropsy evidence of detection bias in the diagnosis of lung cancer. *Arch Intern Med.* 1986; 146(9):1695–1698. [PubMed: 3753108]
15. Sone S, Li F, Yang ZG, et al. Results of three-year mass screening programme for lung cancer using mobile low-dose spiral computed tomography scanner. *Br J Cancer.* 2001; 84(1):25–32. [PubMed: 11139308]
16. Sagawa M, Tsubono Y, Saito Y, et al. A case-control study for evaluating the efficacy of mass screening program for lung cancer in Miyagi Prefecture, Japan. *Cancer.* 2001; 92(3):588–594. [PubMed: 11505403]
17. Veronesi G, Maisonneuve P, Bellomi M, et al. Estimating overdiagnosis in low-dose computed tomography screening for lung cancer: a cohort study. *Ann Intern Med.* 2012; 157(11):776–784. [PubMed: 23208167]
18. Lindell RM, Hartman TE, Swensen SJ, Jett JR, Midthun DE, Mandrekar JN. 5-year lung cancer screening experience: growth curves of 18 lung cancers compared to histologic type, CT attenuation, stage, survival, and size. *Chest.* 2009; 136(6):1586–1595. [PubMed: 19581354]
19. Travis WD, Brambilla E, Noguchi M, et al. International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society international multidisciplinary classification of lung adenocarcinoma. *J Thorac Oncol.* 2011; 6(2):244–285. [PubMed: 21252716]

Table 1

Number of Cancers per Year by Histologic Analysis (Low-Dose Computed Tomography [LDCT] and Chest Radiography [CXR] Arms)

Histologic Subtype	Study Year Associated With First Confirmed Lung Cancer, No. (%)										
	T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	T ₆	T ₇	Subtotal T _{3-T7}	Total T _{0-T7}	
LDCT											
All NSCLC, including BAC and NOS	273 (91.9)	165 (88.7)	204 (86.1)	642 (89.2)	47 (67.1)	83 (77.6)	106 (79.7)	47 (81.0)	1 (100.0)	284 (77.0)	926 (85.0)
BAC only	38 (12.8)	30 (16.1)	29 (12.2)	97 (13.5)	1 (1.4)	5 (4.7)	6 (4.5)	2 (3.4)	0	14 (3.8)	111 (10.2)
Small cell	20 (6.7)	21 (11.3)	27 (11.4)	68 (9.4)	20 (28.6)	21 (19.6)	26 (19.5)	8 (13.8)	0	75 (20.3)	143 (13.1)
Carcinoid	2 (0.7)	0	3 (1.3)	5 (0.7)	1 (1.4)	0	0	0	0	1 (0.3)	6 (0.6)
Unknown	2 (0.7)	0	3 (1.3)	5 (0.7)	2 (2.9)	3 (2.8)	1 (0.8)	3 (5.2)	0	9 (2.4)	14 (1.3)
Total	297 (100.0)	186 (100.0)	237 (100.0)	720 (100.0)	70 (100.0)	107 (100.0)	133 (100.0)	58 (100.0)	1 (100.0)	369 (100.0)	1089 (100.0)
CXR											
All NSCLC, including BAC and NOS	164 (85.0)	110 (82.7)	118 (81.4)	392 (83.2)	92 (80.0)	123 (86.0)	131 (78.9)	53 (73.6)	2 (100.0)	401 (80.5)	793 (81.8)
BAC only	8 (4.1)	3 (2.3)	7 (4.8)	18 (3.8)	5 (4.3)	3 (2.1)	10 (6.0)	0	0	18 (3.6)	36 (3.7)
Small cell	27 (14.0)	20 (15.0)	25 (17.2)	72 (15.3)	22 (19.1)	19 (13.3)	34 (20.5)	16 (22.2)	0	91 (18.3)	163 (16.8)
Carcinoid	1 (0.5)	1 (0.8)	0	2 (0.4)	0	0	0	1 (1.4)	0	1 (0.2)	3 (0.3)
Unknown	1 (0.5)	2 (1.5)	2 (1.4)	5 (1.1)	1 (0.9)	1 (0.7)	1 (0.6)	2 (2.8)	0	5 (1.0)	10 (1.0)
Total	193 (100.0)	133 (100.0)	145 (100.0)	471 (100.0)	115 (100.0)	143 (100.0)	166 (100.0)	72 (100.0)	2 (100.0)	498 (100.0)	969 (100.0)

Abbreviations: BAC, bronchioloalveolar cell carcinoma; NOS, not otherwise specified; NSCLC, non-small cell lung cancer. T indicates study year, beginning with T₀ at baseline.

Table 2

Lung Cancer Counts Used to Derive Overdiagnosis Rates

Lung Cancer Type	LDCT			CXR		
	Not Screen Detected	Screen Detected	Total	Not Screen Detected	Screen Detected	Total
All lung cancers	440	649	1089	690	279	969
All NSCLC, including BAC and NOS	335	591	926	546	247	793
All NSCLC, excluding BAC and including NOS	319	496	815	523	234	757
BAC only	16	95	111	23	13	36

Abbreviations: BAC, bronchioloalveolar cell carcinoma; CXR, chest radiography; LDCT, low-dose computed tomography; NOS, not otherwise specified; NSCLC, non-small cell lung cancer.

Table 3Estimates of P_A and P_S

Lung Cancer Type	Overdiagnosis, % (95% CI)	
	P_A	P_S
All lung cancers	11.0 (3.2 to 18.2)	18.5 (5.4 to 30.6)
All NSCLC, including BAC and NOS	14.4 (6.1 to 21.8)	22.5 (9.7 to 34.3)
All NSCLC, excluding BAC and including NOS	7.1 (−2.3 to 15.6)	11.7 (−3.7 to 25.6)
BAC only	67.6 (53.5 to 78.5)	78.9 (62.2 to 93.5)

Abbreviations: BAC, bronchioloalveolar cell carcinoma; NOS, not otherwise specified; NSCLC, non-small cell lung cancer.

Table 4

Convolution Model Parameter Estimates

Parameter	Estimate (95% CI)	
	NSCLC Excluding BAC	BAC
Sojourn time, ^a mean, y	3.6 (3.0–4.3)	32.1 (17.3–270.7)
Proportion becoming clinical, ^b %		
Within 1 y	24.1 (21–29)	3.1 (0.4–6)
Within 2 y	42.4 (37–49)	6.0 (1–11)
Within 5 y	74.8 (69–81)	14.4 (2–25)
Within 10 y	93.7 (90–97)	26.8 (4–44)
Sensitivity, %		
LDCT	83 (72–94)	38 (7–62)
CXR	33 (26–42)	4 (1–9)

Abbreviations: BAC, bronchioloalveolar cell carcinoma; CXR, chest radiography; LDCT, low-dose computed tomography; NSCLC, non-small cell lung cancer.

^aReciprocal of transition to clinical phase rate parameter; equals mean sojourn time if no competing mortality.

^bPercentage of cases becoming clinical in a given time interval from the start of the preclinical phase (under no screening).

Table 5

Overdiagnosis Rates (P_5) Associated With Low-Dose Computed Tomography (LDCT) Screening Under Various Scenarios

Scenario	P_5^a % (95% CI)					
	NSCLC Excluding BAC		BAC		All NSCLC	
	vs No Screening	vs CXR	vs No Screening	vs CXR	vs No Screening	vs CXR
3 Annual screens						
With 7 y follow-up	21 (16–25)	9 (6–12)	85 (69–93)	71 (52–83)	31 (27–34)	19 (16–23)
With lifetime follow-up	2.6 (2.0–3.3)	1.2 (0.7–1.7)	49 (34–71)	41 (28–62)	11 (7–15)	9 (5–13)
5 Annual screens						
With 5 y follow-up	45 (39–49)	16 (12–21)	91 (84–96)	73 (56–85)	53 (48–56)	25 (21–29)
With lifetime follow-up	3.2 (2.6–3.8)	1.3 (0.8–1.9)	51 (33–71)	42 (27–60)	12 (7–15)	9 (5–13)

Abbreviations: BAC, bronchioloalveolar cell carcinoma; CXR, chest radiography; LDCT, low-dose computed tomography; NSCLC, non-small cell lung cancer.

^aOverdiagnosis rate is total cancers in LDCT cohort minus total cancers in CXR or no screening cohort divided by total number of screen-detected cancers in LDCT cohort.