HHS Public Access

Author manuscript

J Thorac Oncol. Author manuscript; available in PMC 2020 October 01.

Published in final edited form as:

J Thorac Oncol. 2019 October ; 14(10): 1732–1742. doi:10.1016/j.jtho.2019.05.044.

Lung Cancer Incidence and Mortality with Extended Follow-up in the National Lung Screening Trial National Lung Screening Trial Writing Team

Denise R. Aberle,
UCLA Department of. Radiological Sciences

William C. Black,
Dartmouth School of Medicine

Caroline Chiles,
Wake Forest University

Timothy R. Church,
University of Minnesota

Ilana F. Gareen,
Brown University

David S. Gierada,
Washington University School of Medicine

Irene Mahon,
ECOG-ACRIN

Eric A. Miller,
National Cancer Institute

Paul F. Pinsky,
National Cancer Institute

JoRean D. Sicks
Brown University

Abstract

Background—The National Lung Screening Trial (NLST) randomized high-risk current and former smokers to 3 annual screens with either low-dose computed tomography (LDCT) or chest radiographs (CXR) and demonstrated a significant reduction in lung cancer mortality in the LDCT arm after median 6.5 years follow-up. We report on extended follow-up of NLST subjects.

Corresponding author: Paul Pinsky, National Cancer Institute, 9609 Medical Center Dr., Bethesda, MD 20892, UNITED STATES, 240-276-7014, pp4f@nih.gov.

Methods—Subjects were followed by linkage to state cancer registries and the National Death Index. The number needed to screen (NNS) to prevent one lung cancer death was computed as the reciprocal of the difference in the proportion dying of lung cancer across arms. Lung cancer mortality rate ratios (RRs) were computed overall and adjusted for dilution effect, the latter where only deaths with corresponding diagnosis close enough to the end of protocol screening were included.

Results—Median follow-up was 11.3 years for incidence and 12.3 years for mortality. 1701 and 1681 lung cancers were diagnosed in the LDCT and CXR arms, respectively; RR=1.01 (95% CI: 0.95-1.09). Observed lung cancer deaths were 1147 (LDCT) versus 1236 (CXR), RR=0.92 (95% CI: 0.85-1.00). The difference across arms in the number (per 1,000) dying of lung cancer was 3.3, translating into a NNS of 303, similar to the original NNS estimate of around 320. The dilution-adjusted lung cancer mortality RR was 0.89 (95% CI: 0.80-0.997). For overall mortality, there were 5253 (LDCT) and 5366 (CXR) deaths, for a difference across arms (per 1,000) of 4.2 (95% CI: −2.6-10.9).

Conclusion—Extended follow-up of the NLST showed a similar NNS as the original analysis. There was no overall increase in lung cancer incidence in the LDCT versus CXR arm.

Introduction

Lung cancer is the leading cause of cancer death worldwide [1]. Early detection and treatment through screening with low-dose computed tomography (LDCT) has been investigated as a potential means of reducing lung cancer deaths for more than two decades [2–3]. In 2011, a large U.S. study, the randomized National Lung Screening Trial (NLST), reported a significant 20% reduction in lung cancer mortality in high-risk current and former smokers screened annually (3 times) with LDCT as compared to chest radiographs [4,5]. Other small randomized trials, primarily in Europe, have reported mixed results in terms of a lung cancer mortality reduction but were substantially underpowered [6–10]. Recently, the other large LDCT screening trial, NELSON in Europe, reported preliminarily on its findings. Through a 10-year study follow-up period, and following 4 rounds of LDCT screening, NELSON reported a 26% reduction in lung cancer mortality in men (risk ratio=0.74, 95% CI: 0.60-0.91) and a 39% reduction in women (risk ratio=0.61, 95% CI: 0.35-1.04) in the LDCT versus control (non-screening) arm [11].

The median follow-up in NLST as originally reported was 6.5 years, or about 4.5 years following the final scheduled screen [5]. Following the original trial report, an extended follow-up study of the NLST cohort was undertaken, utilizing passive linkages to state cancer registries and the National Death Index (NDI). An additional 5 years of data are now available for lung cancer incidence, and an additional 6 years available for mortality.

The primary objective of the NLST extended follow-up study was to ascertain whether the originally reported reduction in lung cancer mortality in the LDCT versus CXR arm was maintained. With follow-up of 4-5 years following the final screen in the original report, it is possible that earlier detection with LDCT only delayed lung cancer death instead of preventing it. With now 6 additional years of mortality follow-up, it can be observed whether lung cancer deaths were in fact prevented by LDCT screening (at least for a decade) rather

than merely delayed. A secondary objective of extended follow-up was to further assess overdiagnosis in the trial. A modest but statistically significant increase in lung cancer incidence in the LDCT arm, possibly signaling overdiagnosis, was observed with the original follow-up period [12]. With longer follow-up, it is of interest to see whether this increase is preserved.

There are potential issues, however, with examining the extended follow-up data for lung cancer mortality. With follow-up now well beyond the period of trial screening, there is the potential, or even likelihood, of some dilution of the screening effect [13–16]. Specifically, patients in whom cancer did not develop until after the last scheduled screen could not have benefited from the trial screenings; therefore, deaths in such patients would only serve to add noise to the estimates, roughly an equal number of deaths in each arm. Therefore, in analyzing these data, we employ various methods that attempt to control for a dilution effect, including examining the difference across arms in lung cancer deaths in addition to the rate ratio, and examining the rate ratio adjusted for dilution by considering time of diagnosis [13–15]. This latter method, which is well known in the mammography screening trial literature, only includes those cancer deaths for which the corresponding time of cancer diagnosis is close enough to the end of protocol screening in the trial.

Methods

A more detailed description of the NLST has been published previously [4]. Briefly, men and women aged 55-74 years with a minimum of 30 pack-years of cigarette smoking and who were either current smokers or had quit within the past 15 years were enrolled from 2002 to 2004 at 33 medical institutions across the United States. Exclusion criteria included previous lung cancer diagnosis, a CT scan in the prior 18 months, unexplained weight loss in the year before enrollment, or hemoptysis. Participants were randomized into a LDCT or single-view chest radiograph (CXR) arm, with 3 annual protocol screens for each modality.

Participants were actively followed for lung cancer incidence and all-cause mortality until December 31, 2009. During this time, medical records were abstracted for those with a positive screening test or lung cancer diagnosis. Vital status was assessed through annual or semiannual questionnaires and by linkage with the National Death Index (NDI). Institutional review boards at each center approved the study and each person provided written consent to participate in the study.

After the active follow-up period, participants were followed only passively through linkages with state cancer registries and the NDI. Linkages were performed by each participating registry and the NDI using probabilistic linkage methods. Linkages were conducted with cancer registries in the state of the screening center (center's "home state" registry) as well as some neighboring states. For logistical reasons, not all home state registries participated in the linkage effort. In addition, some screening centers did not have participants' names available for linkage purposes, which precluded performing registry linkage for some registries. All centers but one were able to link with NDI. The personally identifiable information NLST had available for linkage included Social Security Number, full name (for some screening centers), date of birth, and sex.

For centers with home state cancer registry linkage (22 of 33, comprising 87.6% of trial participants), lung cancer incidence follow-up was through the end of 2014; otherwise, it was through the end of 2009. Mortality follow-up was through the end of 2015 for centers with NDI linkage (comprising 97.8% of trial participants) and through the end of 2009 for the one center without NDI linkage. See Appendix (Supplemental Table 1) for a summary of linkage efforts by screening center. For assessing mortality from lung cancer, deaths in the original analysis period were evaluated by a death review panel [4]. For the current analysis, the death panel classification was used for those deaths, while the underlying cause of death from the NDI linkage was used for subsequent deaths.

Quantitative Methods

Rates (lung cancer incidence, lung cancer mortality, all-cause mortality) were calculated as the number of events divided by the corresponding person-time; rate-ratios (RRs) were computed as the LDCT arm rate divided by the CXR arm rate. Person time for incidence ended at the end of incidence follow-up, date of lung cancer diagnosis or date of death, whichever came first. Person time for mortality ended at the end of mortality follow-up or death, whichever came first. In addition to rates and RRs, for each event type we computed the proportion of subjects in each arm with the event and the difference across arms in those proportions. Note that, unlike the RR, the expected difference in proportions is not affected by dilution, since an equal number of events in each arm occurring beyond the time where screening could have an effect cancel each other out on average. The number needed to screen (NNS) to prevent one lung cancer death was calculated as the reciprocal of the difference across arms in the proportion dying of lung cancer. Potential interactions of several risk factors with trial arm, specifically age, sex and smoking status (current versus former smoker), were assessed using Poisson regression. The distribution of lung cancer cases by histology and stage was analyzed using chi-squared tests. The overdiagnosis rate was calculated as the difference across arms in lung cancer cases divided by the number of LDCT screen-detected cases.

Analysis Adjusted for Dilution

To derive the dilution-adjusted lung cancer mortality RR, the cutoff time for cancer diagnosis must first be determined; only those lung cancer deaths (in each arm) for which the corresponding diagnosis is before this cutoff time are included in the RR computation. One proposed method is to assess when in study-time cumulative cancer incidence across arms first becomes equalized [13–14]. If screening results in overdiagnosis, incidence would never become equalized across arms. In NLST, while there was overdiagnosis based on the original data, the majority of overdiagnosed cases were identifiable by histology [12]. Almost all cases classified as bronchioloalveolar carcinoma (BAC) were overdiagnosed, and BAC represented the majority of all overdiagnosed cases [12]. Therefore, to define the cutoff time for the dilution-adjusted analysis, incidence across arms was examined by study year, excluding BAC cases, and the cutoff time was defined as the (end of) the earliest study year for which there was no significant difference in cumulative incidence across arms. As a sensitivity analysis, we also examined dilution-adjusted RRs using alternative study year cutoff times.

Analyses using Calendar Time versus Study Time

Lung cancer mortality results for NLST based on the original data were first reported using a January 15th, 2009 cutoff date in accordance with the interim analysis plan and to account for time lags associated with the endpoint verification process; a December 31, 2009 cutoff was used for all-cause mortality [5]. Lung cancer mortality results were subsequently reported using all events through the later cutoff date (December 31, 2009) [17]. Because subjects were enrolled in NLST over roughly a two-year period, these calendar time cutoffs resulted in a range of times on study for the original analysis, with median (interquartile range) of 5.5(5.2-5.9) and 6.5 (6.1-6.9) years for the earlier and later dates, respectively. From a scientific standpoint, analyses based on study time cutoffs, where all subjects have essentially the same time on study, are more meaningful since they allow assessment of all events within a given time after randomization and protocol screens. The extended follow-up data allow us now to compute lung cancer mortality RRs based on study time cutoffs with similar median follow-up times as those in the original analyses.

Results

A total of 26,722 and 26,730 participants were randomized to the LDCT and CXR arms, respectively. Baseline participant demographics and smoking history were similar across arms (Table 1).

Median follow-up time for incidence and mortality was similar across arms. For incidence, median (25th/75th) follow-up was 11.3 (9.0/11.8) years in the LDCT arm and 11.3 (8.9/11.8) years in the CXR arm; for mortality, median (25th/75th) follow-up was 12.3 (11.9/12.8) years in each arm.

Lung Cancer Incidence

Figure 1 (A,B) shows cumulative lung cancer incidence by arm. There were 1701 lung cancer cases in the LDCT arm versus 1681 in the CXR arm, giving a RR of 1.01 (95% CI: 0.95-1.09) (Figure 1A). For all cases excluding BAC, the RR was slightly under one (RR=0.97, 95% CI: 0.90-1.04), whereas there was a significant increase in BAC cases in the LDCT arm (RR=2.6, 95% CI: 1.9-3.7) (Figure 1B). As seen in Figure 1A, the excess cumulative number of cases in the LDCT versus CXR arm peaks around year 3, the end of the screening phase of the trial, and declines thereafter. Overall lung cancer rates per 10,000 PY were 63.8 and 62.9 in the LDCT and CXR arms, respectively. The overdiagnosis rate was 3.1% (20/649) overall and 79% (75/95) for BAC.

Lung Cancer Characteristics

Table 2 shows the distribution of histology and stage by arm. With the exception of BAC, the histology distribution was generally similar across arms. In terms of stage, a significantly higher proportion of LDCT versus CXR arm cases were stage I, 39.6% versus 27.5%, $p < 0.0001$ (excluding BAC, the stage I proportions were 37% versus 27%, $p < 0.0001$). Conversely, a significantly lower proportion of cases in the LDCT versus CXR arm were stage IV, 27.5% versus 35.5% ($p < 0.0001$).

Lung Cancer Mortality, Stage IV Disease, and All-Cause Mortality

Table 3 shows lung cancer mortality rates across arms. There were 1147 deaths (42.9 per 1,000 subjects) from lung cancer in the LDCT arm versus 1236 (46.2 per 1,000) in the CXR arm. The difference across arms (CXR minus LDCT) in the number of subjects (per 1,000) dying of lung cancer was 3.3 (95% CI: −0.2-6.8; p=0.06), which translates into a NNS of 303. The RR for lung cancer mortality was 0.92 (95% CI: 0.85-1.00; p=0.05). The lung cancer mortality RR was lower for women (RR=0.86) than for men (RR=0.97), lower for current (RR=0.88) than for former smokers (R=1.01) and lower for subjects 55-64 at entry (RR=0.86) than those 65-74 (RR=1.01) (Table 3). However, the interactions of trial arm by sex and by smoking status were not statistically significant, indicating there was no statistical difference in the RRs by sex or smoking status. The interaction of trial arm by age was borderline significant (p=0.051).

For the analysis adjusted for dilution, the cumulative incidence RR across arms (excluding BAC) first became non-significant at study year 6, with a RR=1.07 (p=0.13). Therefore, for the dilution-adjusted analysis, only those deaths with diagnosis through study year 6 were included. There were 578 such lung cancer deaths in the LDCT arm versus 646 in the CXR arm, giving a lung cancer mortality RR of 0.89 (95% CI: 0.80-0.997; p=0.043) (Table 3). The difference in the number across arms dying of lung cancer (per 1,000) based on the dilution-adjusted analysis was 2.5 (95% CI: 0.001-5.1; p=0.05), giving a NNS of 394.A similar pattern was observed as in the overall analysis of the RR being lower in women (RR=0.80) than men (RR=0.95), current (RR=0.84) than former smokers (RR=0.99) and younger (RR=0.85) versus older (RR=0.94) subjects, though none of these interactions with trial arm were statistically significant. Supplemental Table 2 shows cumulative incidence RRs for various alternative study time cutoffs and the corresponding dilution-adjusted RRs.

There were 468 stage IV cases in the LDCT arm versus 597 in the CXR arm, giving a RR of 0.79 (95% CI: 0.70-0.89) (Table 3). Considering the same study period as the dilution-adjusted analysis (through study year 6), there were 245 (LDCT) versus 344 (CXR) cases, RR=0.72 (95% CI:0.61-0.84). There were no significant interactions by sex, age or smoking status, either for all stage IV cases or for stage IV cases through year 6.

Figures 2A and B show lung cancer deaths over time for the overall and dilution-adjusted analyses, as well as stage IV cases over time.

Table 4 shows lung cancer mortality RRs for comparable time periods for the originally reported and extended follow-up data, based on calendar time and study time cutoffs, respectively. For similar median follow-up time, RRs were similar. For example, at median 5.5 years follow-up for both the calendar and study time cutoffs, RRs were 0.80 and 0.81, respectively (Table 4). Going from 6 to 7 study years, however, the RR increased substantially, from 0.81 to 0.86, a result of the greater number of lung cancer deaths in the LDCT than CXR arm in study year 7 (see Figure 2A).

Overall mortality results by arm are shown in Table 3. The overall mortality RR was 0.97 (95% CI: 0.94-1.01), with a difference across arms in the number dying (per 1,000) of 4.2

(95% CI: −2.6 – 10.9; p=0.18). The distribution of causes of death was similar across arms (Supplemental Table 3).

Discussion

In this extended follow-up analysis of the NLST, the difference in the proportion dying of lung cancer across arms (CXR minus LDCT) was 3.3 per 1,000, which translates into a NNS to prevent one lung cancer death of 303. This 3.3 per 1,000 difference was similar to that observed in prior analyses of the original trial data, based either on the Jan 15th, 2009 cutoff (3.2 per 1,000) or December 31, 2009 cutoff (3.1 per 1,000), and the NNS of 303 was similar to earlier reported NNS values of around 320 [5,17]. The stability of this difference over time indicates that LDCT screening did not just delay lung cancer death by a few years, but prevented it, or at least delayed it for more than a decade.

In contrast to the stability over time of the difference in lung cancer deaths across arms, and by extension the NNS, the rate ratio (RR) for lung cancer mortality changed substantially over time. The RR derived from the original data increased from 0.80 to 0.84 based on a relatively small (about one year) difference in the calendar-time cutoff. With extended follow-up well beyond the end of protocol screening, the RR would be expected to move towards the null due to dilution of the screening effect, and this was in fact observed, with an RR of 0.92. However, for the dilution-adjusted analysis, the RR was 0.89, showing a smaller mortality reduction than earlier analyses. While this RR was adjusted for dilution, dilution still may have affected the estimate, as the 4-year post-screening window for diagnosis likely included some cancers with short lead times whose outcome could not have been affected by screening. Although mortality RR estimates from trials are an important public health tool for assessing screening benefits, they are problematic because a standard screening trial, with several rounds of screening and some additional years of follow-up, does not match up exactly with screening as performed in the population setting. As seen here, small changes in follow-up time can lead to non-trivial changes in RR. With the original NLST findings, modeling efforts attempted to extrapolate trial results to the population screening setting [18,19]. Additional modeling efforts incorporating these extended follow-up data, and the results of the NELSON trial, may prove useful for informing both the population and individual perspectives, the latter of which is most appropriate for shared-decision making.

The p-values for the lung cancer mortality RR and difference in proportions hovered around the 0.05 level. P-values were not emphasized because the null hypothesis of no lung cancer mortality difference across arms has already been rejected by the original analysis. The dilution effect, of adding (roughly) equal numbers of events in each arm, in addition to moving the RR, but not the difference in proportions, towards the null, also increases the standard deviation (SD) of both the RR and the difference in proportions and thus tends to increase the associated p-value. For example, counting only lung cancer deaths occurring within 6.5 years of randomization, there were 457 and 550 in the LDCT and CXR arms, respectively, giving a difference in proportions of 3.5 (per 1,000) and a corresponding SD of 1.1. For total follow-up, approximately equal numbers of deaths were added (690 and 686 in the LDCT and CXR arms, respectively), giving a similar difference in proportions (3.3) but a substantially inflated SD of 1.8, which caused the p-value to increase from 0.003 to 0.06.

The additional nearly 700 deaths in each arm also caused the RR to increase towards the null, from 0.83 to 0.92, and the p-value to increase from 0.003 to 0.05.

The reduction in stage IV disease across arms was greater than that of lung cancer deaths. Since stage IV cases have a high case-fatality rate, the difference across arms in deaths from stage IV cancers (N=140) was similar to the difference across arms in overall stage IV cases (N=129). However, the difference in lung cancer deaths across arms was only 89, because the 140 fewer deaths from stage IV cancers in the LDCT arm were partially offset by an excess in the LDCT arm of 31 deaths from stage I-III and 20 deaths from unknown stage cancers. Therefore, some of the difference across arms in stage IV cases may have been the result of earlier diagnosis in the LDCT arm, at a time when metastases were not clinically apparent, of tumors that eventually progressed, in spite of early diagnosis and treatment, to metastatic disease.

In contrast to what was observed with the original follow-up, in this extended follow-up analysis there was no statistically significant reduction in all-cause mortality in the LDCT versus CXR arm. However, as described above, for the same difference in proportions, the p-value is substantially higher in the extended follow-up compared to original analysis due to the extra noise associated with the dilution effect. For all-cause mortality, the difference across arms in the proportion dying was 4.6 per 1,000 in the original analysis and a similar 4.2 per 1,000 here, indicating that the all-cause mortality difference was essentially sustained. The p-value, though, increased from 0.02 to a non-significant 0.18 in the extended follow-up. Therefore, the current lack of a statistically significant effect for all-cause mortality should not be taken to negate the original significant finding; it is more likely related to using the “incorrect window” for follow-up (i.e., too long a period post-screening) [16] . In addition, with respect to non-lung cancer mortality, the original RR was 0.96 with a corresponding non-significant p-value of 0.29; therefore, this is not inconsistent with the currently observed non-significant RR of 0.99 for non-lung cancer mortality.

As with the originally reported results, in the updated analysis there was an observed lower RR for lung cancer mortality (i.e., greater percentage mortality reduction with LDCT) in women than in men, although the interaction of gender and trial arm was not statistically significant. Preliminary results from the NELSON trial also show a greater observed percentage mortality reduction in women, although it is not clear whether this represents a statistically significant difference [11]. There were also some observed differences here in lung cancer mortality RRs by age and smoking status, but given the non-significant interaction p-values, it is not clear whether these are real. Note for age, the p-value was borderline significant (0.05) for the overall analysis but not close to significant (p=0.39) for the dilution-adjusted analysis, and also that the analysis of interactions involved multiple comparisons. For gender, as well as age and smoking status, meta-analyses of all LDCT trials may shed some light on whether the effect of LDCT screening is truly differential by these factors. From a public health standpoint, even if the RRs were the same according to, say, smoking status, the higher background lung cancer rate for current versus former smokers indicates that the risk difference (difference in proportions across arms dying of lung cancer) would be greater, and correspondingly, that the NNS would be lower, in current versus former smokers.

Additionally, NLST was not powered for interactions, so modest, but potentially clinically significant, interactions of the RR with the factors of age, sex or smoking status could have failed to reach statistical significance. However, within this trial population, which was all at high risk due to smoking history but also generally healthy, it is unclear what the biological rationale would be for an interaction with these factors; thus, any true interactions would likely be of small magnitude. With more varied populations potentially undergoing LDCT screening, this might not be the case, as factors related to ability to undergo curative treatment or to differential lung cancer histology might alter the effectiveness of LDCT.

After 11.3 median years follow-up for incidence, or 9.3 years after the last scheduled screen, lung cancer incidence was similar across arms, with a RR of 1.01 (95% CI: 0.95-1.09) for the LDCT versus CXR arm. In contrast, in the original trial period of 6.5 years median follow-up, there was a significantly elevated RR of 1.13 [5]. This indicates that so-called "catch-up" likely occurred in the CXR arm, where the counterparts of those cancers diagnosed early in the LDCT arm were eventually diagnosed in the CXR arm. A mathematical model of lung cancer natural history fit to the original NLST data predicted that 94% of cases, excluding BAC, would become clinically apparent within 10 years of LDCT screen diagnosis [12]. Since the average follow-up of LDCT screen-diagnosed cases is now about 10 years, the 94% estimate is generally consistent with the current observation of no increase in (non-BAC) lung cancer in the LDCT arm. In contrast, there continued to be a large excess of BAC cases in the LDCT (N=121) versus CXR arm (N=46), with few additional cases identified after the original follow-up period. This is also consistent with the above-mentioned model's predictions, which estimated that only around 25% of screen detected BAC would become clinically apparent within 10 years. Some BAC cases could eventually present clinically after more than 10 years, so the 79% overdiagnosis estimate for BAC could be an overestimate.

In 2011, a multi-society committee recommended changes to the classification of lung adenocarcinoma, reclassifying BAC into new categories of adenocarcinoma in situ (AIS), minimally invasive adenocarcinoma (MIA), invasive lepidic adenocarcinoma and invasive mucinous adenocarcinoma, and discontinuing use of the term BAC [20,21]. The new categories involve the same ICDO morphology codes as previously used for BAC, with the exception of new codes for MIA. These same codes were used through the entire NLST follow-up period to define BAC (or what was formerly known as BAC); thus, the reclassification should not have affected the overdiagnosis estimate for BAC. Note MIA tumors were not ascertained in NLST.

The magnitude of overdiagnosis as estimated from LDCT screening trials depends critically on the length of follow-up following the final screen. For NLST, the overdiagnosis rate decreased from 18% in the original analysis (median 4.5 years follow-up after the last screen) to 3% with extended follow-up. However, even controlling for follow-up time, there is great variability in overdiagnosis rates across trials. In the Danish trial, after median 5 years of follow-up following the final screen, the overdiagnosis rate was 67%, while in the Italung trial, with median 4.5 years of such follow-up, the overdiagnosis rate was 0 [6,22]. More research is needed concerning the factors that influence overdiagnosis in LDCT screening.

A limitation of the analysis was that use of LDCT screening after the original trial period was not ascertained. NLST participants were sent a letter in 2010 summarizing trial results, with CXR arm subjects told that they may want to discuss LDCT screening with their health-care provider and LDCT arm subjects told they may want to discuss continuing screening. However, LDCT screening was not generally covered by private insurance or Medicare until 2015, and survey evidence suggests that usage was low in the U.S. through 2015 [23,24]. However, as trial volunteer participants, NLST subjects may have been more motivated to receive screening than eligibles in the general population. In addition, indirect evidence suggests there was little LDCT screening among NLST participants following the screening phase of the trial. As described above, there were few cases of BAC (or in the new terminology, but with the same morphology codes, invasive lepidic or mucinous adenocarcinoma), following the screening phase of the trial. After 90 cases of BAC in the LDCT arm during the three screening phase years of the trial (T0-T2), including 24 after the 2nd incidence (T2) screen, there were only an average of 4 per year (31 total) for the next 8 years in the LDCT arm, and a similar number during that period in the CXR arm. Since BAC cases are generally only found with LDCT screening, such screening was likely low, and similar across trial arms, in the post-screening phase of the trial. Another limitation was that death review was not performed for deaths following the original analysis period. However, an analysis of the agreement between death certificates and death review for the original period showed high levels of agreement and minimal effect on the lung cancer mortality RR [25].

Conclusion

With further follow-up of NLST subjects, the originally reported reduction in lung cancer deaths in the LDCT versus CXR arm was sustained; in contrast, the originally reported increase in lung cancer incidence was no longer observed.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

Cancer incidence data have been provided by the following state cancer registries: Alabama, Arizona, California, Colorado, District of Columbia, Georgia, Hawaii, Idaho, Indiana, Iowa, Kentucky, Louisiana, Maryland, Massachusetts, Michigan, Minnesota, Missouri, Nevada, North Carolina, Ohio, Pennsylvania, Rhode Island, Texas, Utah, Virginia and Wisconsin. All are supported in part by funds from the Centers for Disease Control and Prevention, National Program for Central Registries, local states, or by the National Cancer Institute, Surveillance, Epidemiology, and End Results Program. The results reported here and the conclusions derived are the sole responsibility of the authors.

DISCLOSURES:

Dr. Aberle reports grants from American College of Radiology Imaging Network (ACRIN), grants from American College of Radiology Imaging Network (ACRIN), during the conduct of the study; personal fees and other from LUNGevity Foundation, other from Siemens Medical Solutions USA, personal fees from 2012: First Annual McLennan Lecture, University of Iowa, personal fees from 2012: US Department of Veteran Affairs, personal fees from 2012: Gordon Gamsu Memorial Lecture, Society of Thoracic Radiology Annual Meeting, personal fees from 2012: University of Texas Southwestern Medical Center Meeting, personal fees from 2012: Vanderbilt-Ingram Cancer Center Retreat on Lung Cancer, personal fees from 2012: ARRS Chest Imaging Symposium, personal fees from 2012: Institute of Medicine, Affordable Care for the 21st Century Workshop, personal fees from 2012:

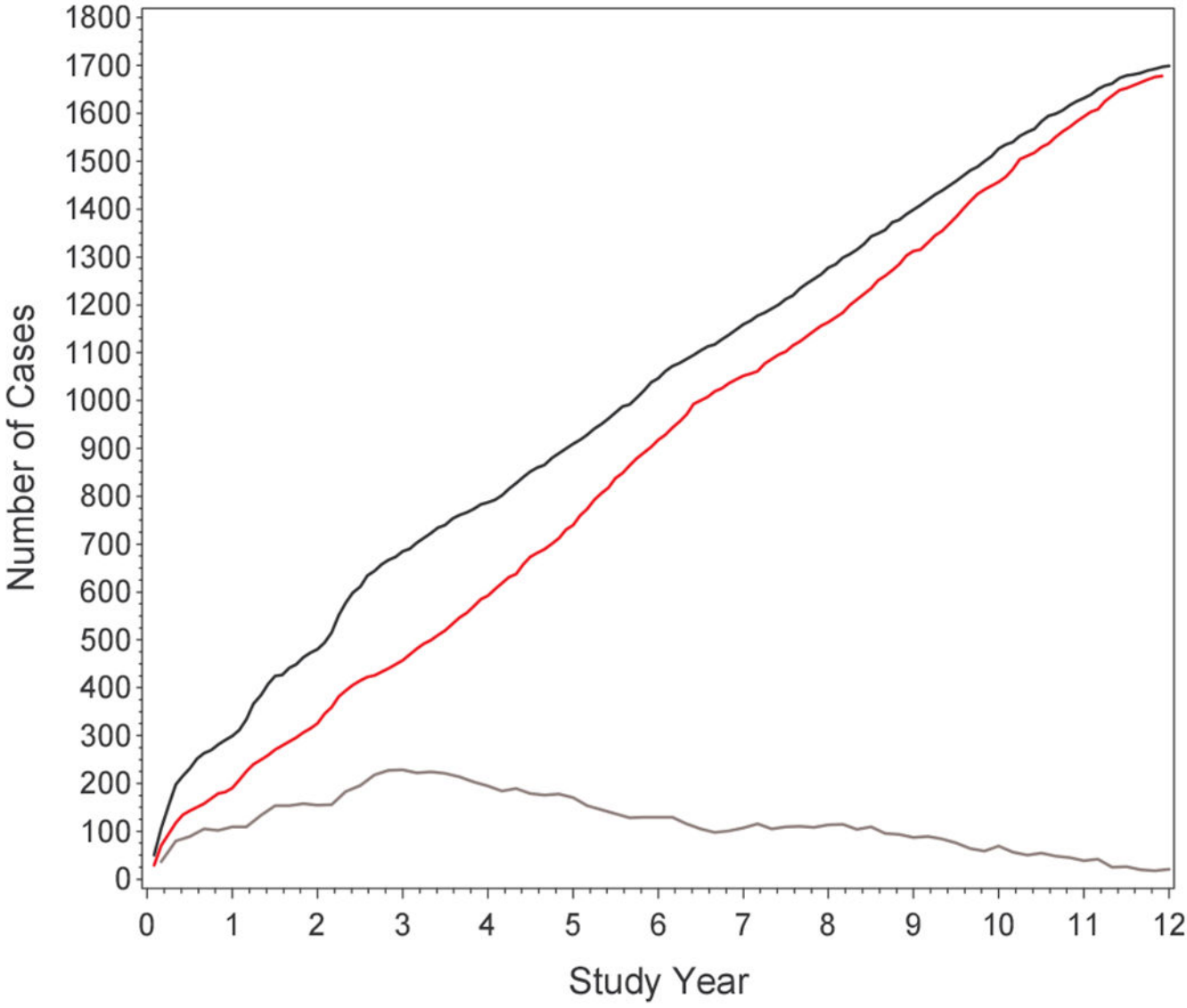
California Technology Assessment Forum Policy Symposium, personal fees from 2012: Stanford University, Department of Radiology Grand Rounds Series, personal fees from 2012: New York University, Head to Toe Imaging Conference, personal fees from 2012: Society of Thoracic Surgeons (STS) 48th Annual Meeting, personal fees from 2012: 13th International Lung Cancer Congress, personal fees from 2012: European Society of Thoracic Imaging 2012 Annual Meeting., personal fees from 2013: Glendale Memorial Hospital and Health Center's Continuing Medical Education, personal fees from 2013: Colorado Radiological Society Visiting Professor Series, personal fees from 2013: University of Colorado, Denver, Grand Rounds Series, personal fees from 2013: 13th International Lung Cancer Congress, personal fees from 2013: International Association for the Study of Lung Cancer 15th World Conference, non-financial support from 2012: National Cancer Advisory Board Meeting, non-financial support from 2012: ACRIN Semi-Annual Meetings, non-financial support from 2012: Eastern Cooperative Oncology Group- American College of Radiology Imaging Network Strategic Retreat, non-financial support from 2012: Eastern Cooperative Oncology Group-American College of Radiology Imaging Network Imaging Workshop, non-financial support from 2012: AACR-IASLC 2nd Annual Joint Conference, non-financial support from 2013: Institute of Medicine Affordable Care Workshop, non-financial support from 2013: 3rd World Congress on Thoracic Imaging (WCTI)., non-financial support from 2013: National Cancer Institute Lung SPORE Workshop, non-financial support from 2013:ACRIN Semi-Annual Meetings, personal fees from 2014: AACR-IASLC 3rd Joint Conference, non-financial support from 2014: European Congress of Radiology, non-financial support from 2014: University of California Dose Retreat, non-financial support from 2014: ECOG-ACRIN Semi-Annual Meetings, personal fees and non-financial support from 2014: American Thoracic Society, non-financial support from 2014: DECAMP Consortium Meeting, non-financial support from 2014: National Cancer Institute Lung SPORE Workshop, non-financial support from 2014: AACR Lung Cancer Screening Meeting, non-financial support from 2014: LUNGevity Science Advisory Board Meeting, personal fees from 2014: LUNGevity Award Program Review, personal fees from 2015: LARS Midwinter Conference, non-financial support from 2015: STR Annual Meeting, non-financial support from 2015: Cambridge Chest Meeting, non-financial support from 2015:ACRIN Semi-Annual Meetings personal fees from 2015: LUNGevity Award Program Review, non-financial support from 2015: AACR Annual Meeting, non-financial support from 2015: DECAMP Consortium Meeting, other from 2016 Veracyte Advisory Board, non-financial support from 2015: NIH SPORE, grants from 2015: Lung Nodule Surveillance Trail (LNST) Kick off Meeting, non-financial support from 2015: Siemens Annual UCLA Research Meeting, non-financial support from 2015: IASLC WCLC Annual Meeting, non-financial support from 2015: LUNGevity Scientific Advisory Board Meeting, grants from 2015: MCL Consortium Meeting, non-financial support from 2015: ECOG-ACRIN Semi Annual Meeting, from 2015: Veracyte, non-financial support from 2016: Moffitt Cancer Center, non-financial support from 2016: Yale University Grand Rounds, non-financial support from 2016: Kaiser Radiology Symposium grants and non-financial support from 2016: MCL Consortium Meeting, non-financial support from 2016: ECOG-ACRIN Semi Annual Meeting, non-financial support from 2016 DECAMP Consortium Meeting, other from 2016: American Society of Clinical Oncology (ASCO) Annual Meeting, non-financial support from 2016: National Academy of Sciences Workshop, non-financial support from 2016: Siemens Annual UCLA Research Meeting, personal fees from 2016: Veracyte Advisory Board Meeting, personal fees and non-financial support from 2016: International Lung Cancer Congress (ILCC) Meeting, grants and non-financial support from 2016: MCL Consortium Meeting, non-financial support from 2016: Harvard University Chest Imaging Course, personal fees from 2016: Torrance Medical Center Grand Rounds, from null, outside the submitted work.

Jolean Sicks reports a grant to the institution from NCI, U10CA18-820-0151, during the study. Caroline Chiles reports a grant to the institution from NIH, CA 80098, during the study. Remaining authors have nothing to disclose.

References

1. World Health Organization. GLOBOCAN 2012: Estimated cancer incidence, mortality and prevalence worldwide in 2012. http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx Assessed Oct 10th, 2018.
2. Naidch DP, Marshall CH, Gribbin C, Arams RS, McCauley DI. Low-dose CT of the lungs: preliminary observations. Radiology 1990; 175: 729–731. [PubMed: 2343122]
3. Kaneko M, Eguchi K, Ohmatsu H, et al. Peripheral lung cancer: screening and detection with low-dose spiral CT versus radiography. Radiology 1996; 201: 798–802 [PubMed: 8939234]
4. Aberle DR, Berg CD, Black WC, et al. The National Lung Screening Trial: overview and study design. Radiology. 1 2011;258(1):243–253. [PubMed: 21045183]
5. The 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]
6. Paci E, Puliti D, Pegna AL, et al. Mortality, survival and incidence rates in the ITALUNG randomized lung cancer screening trial. Thorax 2017; 72: 825–831. [PubMed: 28377492]

7. Infante M, Cavuto S, Lutman FR, et al. Long-term follow-up results of the DANTE Trial, a randomized study of lung cancer screening with spiral computed tomography. Am J Resp Criti Care Med 2015; 191: 1166–1175.
8. Wille MW, Dirksen A, Ashraf H, et al. Results of the randomized Danish Lung Cancer Screening Trial with focus on high-risk profiling. Am J Resp Criti Care Med 2016; 193: 542–551.
9. Pastorino U, Rossi M, Rosato V, et al. Annual or biennial CT screening versus observation in heavy smokers: 5-year results of the MILD trial. Eur J Cancer Prev 2012; 21: 308–315. [PubMed: 22465911]
10. Doroudi M, Pinsky PF, Marcus PM. Lung cancer mortality in the Lung Screening Study (LSS) feasibility trial. JNCI Cancer Spectrum 2018; doi: 10.1093/jncics/pky042.
11. De Koning H. Mortality results of the NELSON randomized controlled population-based screening trial; IASLC 19th World Conference on Lung Cancer; Toronto, Canada: Sept 23–26
12. Patz E, Pinsky PF, Gatsonis C, et al. Overdiagnosis in low-dose computed tomography screening for lung cancer. JAMA Intern Med 2014; 174: 269–274. [PubMed: 24322569]
13. Etzioni R, Self SG. On the catch-up time method for analyzing cancer screening trials. Biometrics 1995; 51: 31–43. [PubMed: 7766785]
14. Prorok PC, Aron JL. An analysis of the mortality effect in a breast cancer screening study. Int J Epidemiol 1986; 15: 36–43. [PubMed: 3957541]
15. Nelson HD, Fu R, Cantor A, et al. Effectiveness of breast cancer screening: systematic review and meta-analysis to update the 2009 U.S. Preventive Services Task Force recommendation. Ann Int Med 2016; 164: 244–255. [PubMed: 26756588]
16. Hanley J Analysis of mortality data from cancer screening studies: looking in the right window. Epidemiology 2005: 16: 786–790. [PubMed: 16222169]
17. Pinsky PF, Church TR, Izmirlian G, Kramer BS. The National Lung Screening Trial: Results stratified by demographics, smoking history and lung cancer histology. Cancer 2013; 119: 3976–3983. [PubMed: 24037918]
18. De Koning H, Meza R, Plevritis SK, 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 Int Med 2014; 160: 311–320. [PubMed: 24379002]
19. McMahon PM, Meza R, Plevritis SK, et al. Comparing benefits from many possible computed tomography lung cancer screening programs: extrapolating from the National Lung Screening Trial using comparative modeling. PLoS One 2014; 9(6): e99978. [PubMed: 24979231]
20. 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 Thoracic Oncol 2011; 6: 244–248.
21. Travis WD, Brambilla E, Nicholson AG, et al. The 2015 World Health Organization Classification of Lung Tumors. J Thorac Oncol 2015; 10: 1243–1260. [PubMed: 26291008]
22. Heleno BN, Siersma V, Brodersen J. Estimation of overdiagnosis of lung cancer in low-dose computed tomography screening: a secondary analysis of the Danish Lung Cancer Screening Trial. JAMA Int Med 2018; 178: 1420–1422.
23. Jemal A, Fedewa SA. Lung cancer screening with low-dose computed tomography in the United States-2010 to 2015. JAMA Oncol 2017; 3: 1278–1281. [PubMed: 28152136]
24. Huo J, Shen C, Volk RJ, Shih Y. Use of CT and chest radiography for lung cancer screening before and after publication of screening guidelines: intended and unintended uptake. JAMA Int Med 2017; 177: 439–441.
25. Marcus PM, Doria-Rose VP, Gareen IF, et al. Did death certificates and a death review process agree on lung cancer cause of death in The National Lung Screening Trial? Clin Trials 2016; 13: 434–438. [PubMed: 27006427]



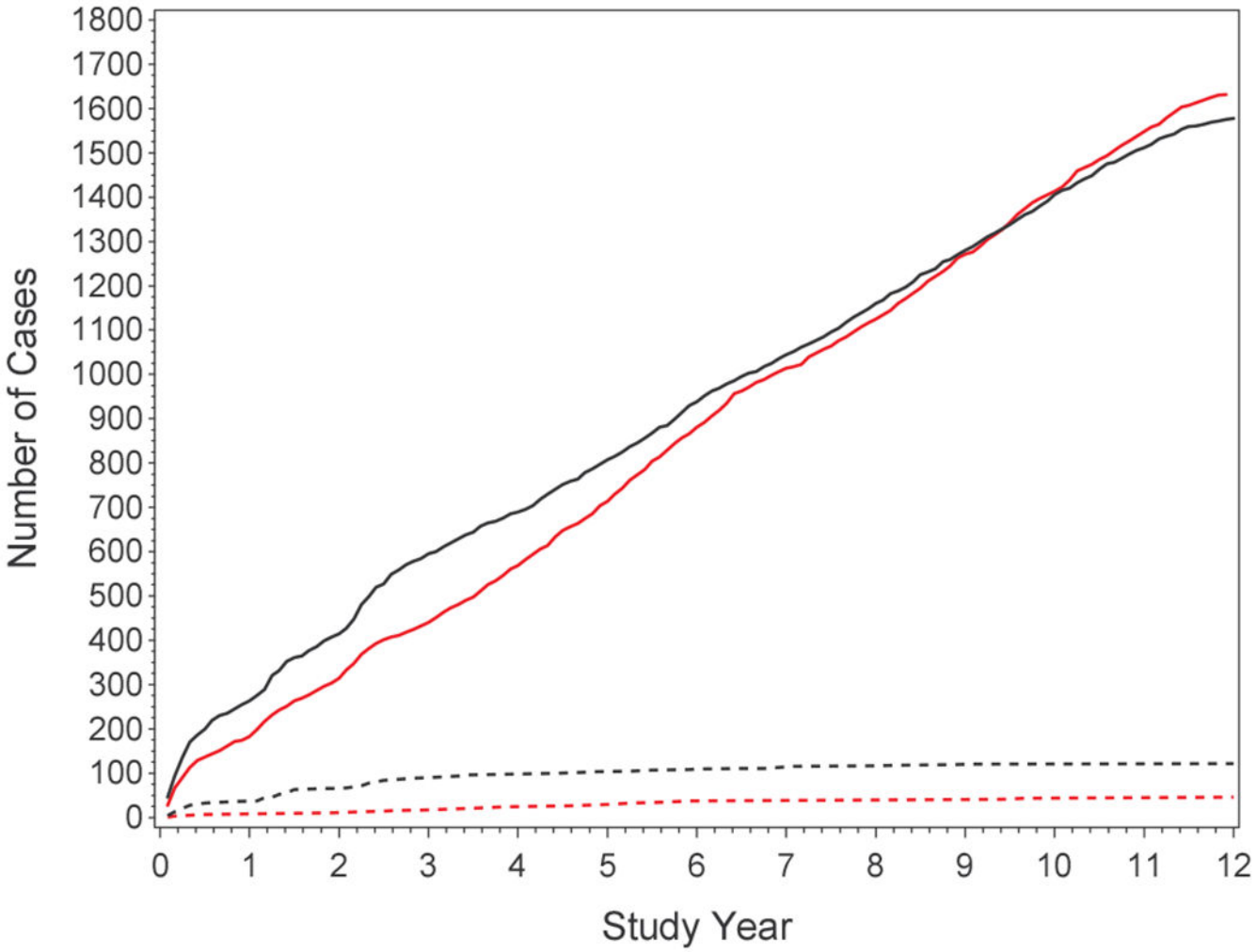
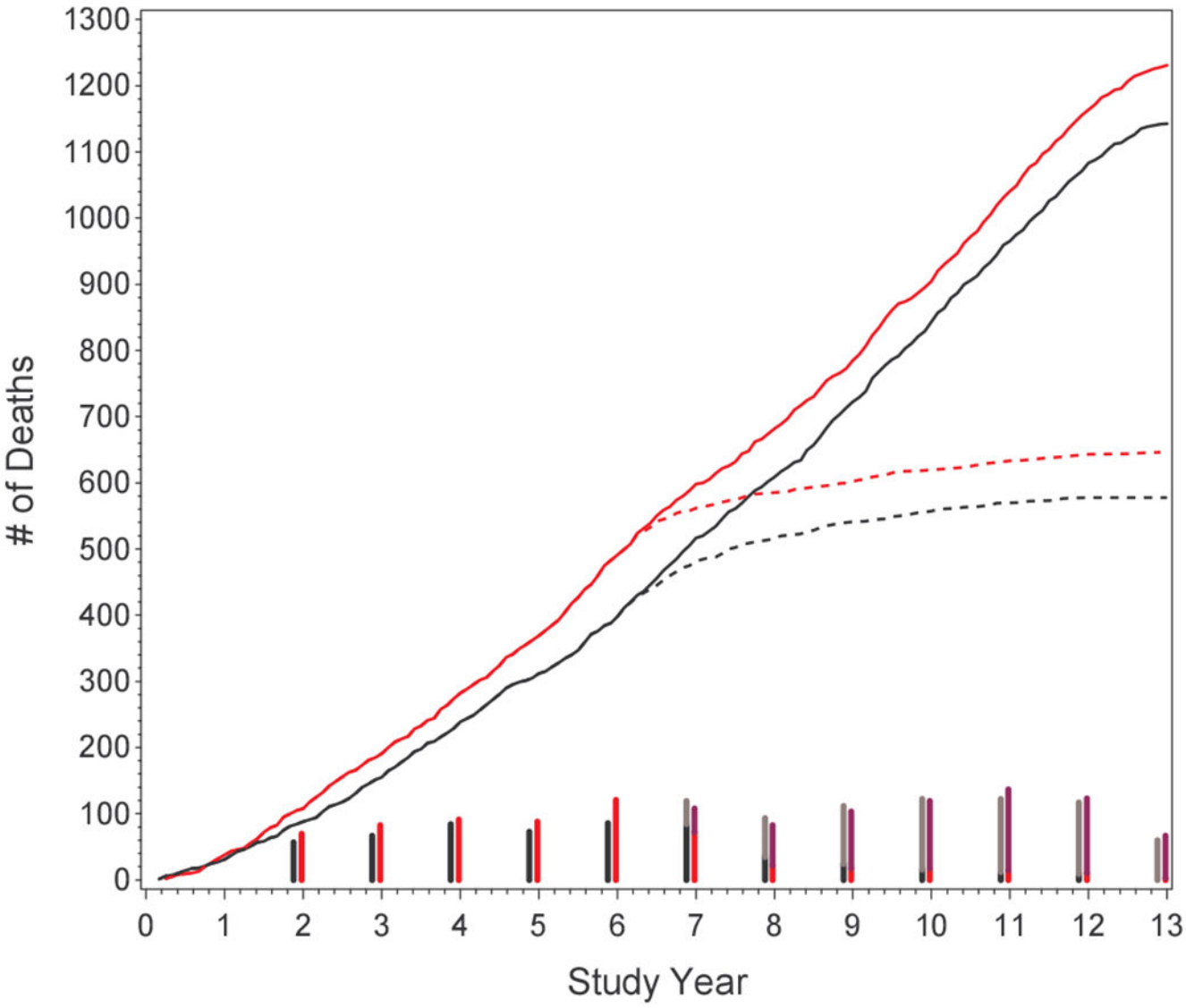


Figure 1.

Cumulative lung cancer cases by arm. A) All lung cancers. Black is LDCT arm, red is CXR arm. Gray line represents excess of cases in the LDCT arm over the CXR arm. B) All cases excluding bronchioloalveolar carcinoma (BAC) are solid lines, BAC cases are dotted lines. Black is LDCT arm, red is CXR arm.



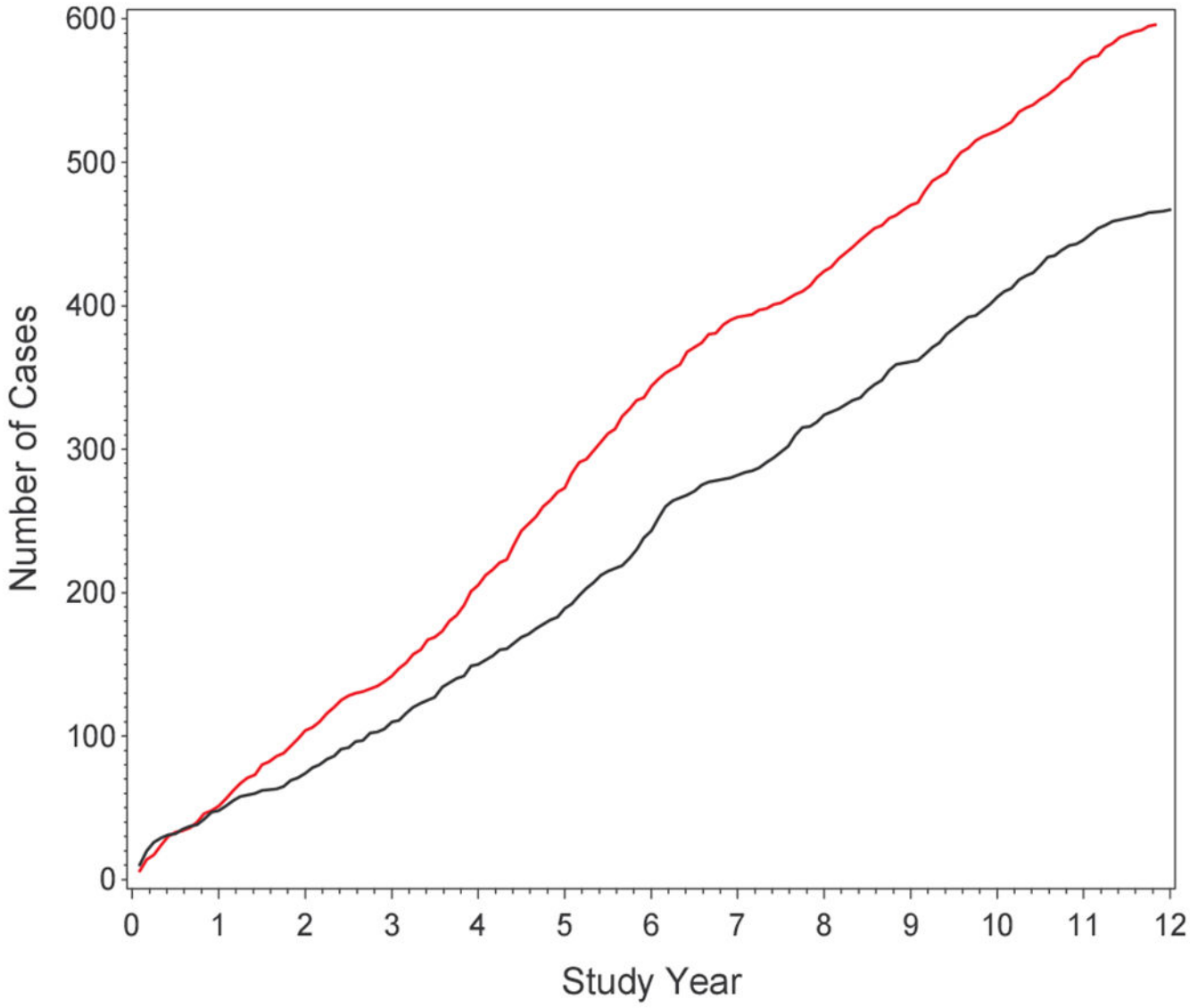


Figure 2.

A. Lung cancer deaths by trial arm. Black is LDCT arm, red is CXR arm; solid lines are all deaths and dotted lines show deaths for the dilution-adjusted analysis. Vertical bars show number of deaths for each study year. Black/gray and red/purple bars are for CXR and LDCT arms, respectively. Total height is all deaths, height of black/red segment shows number of deaths for dilution-adjusted analysis.

B. Stage IV lung cancers cases by trial arm. Black is LDCT arm, red is CXR arm.

Table 1.

Demographics and Smoking History

	LDCT(N=26722)	CXR(N=26730)
Men	15769 (59.0)	15761 (59.0)
Women	10953 (41.0)	10969 (41.0)
Non-Hispanic White	23953 (89.6)	23949 (89.6)
Non-Hispanic Black	1187 (4.4)	1174 (4.4)
Hispanic	479 (1.8)	456 (1.7)
Asian	546 (2.0)	525 (2.0)
American Indian/Native Alaskan	87 (0.4)	97 (0.4)
Native Hawaiian/Pacific Islander	83 (0.3)	81 (0.3)
Other/Unknown	387 (1.5)	448 (1.7)
Current Smoker	12860 (48.1)	12900 (48.3)
Former Smoker	13862 (51.9)	13830 (51.7)
Median (25^{th}/75^{th}) Pack-Years	48 (39/66)	48 (39/66)
Age at Enrollment		
55-59	11442 (42.8)	11423 (42.7)
60-64	8170 (30.6)	8199 (30.7)
65-69	4756 (17.8)	4761 (17.8)
70-74	2354 (8.8)	2347 (8.8)

Table 2.

Histology and Stage of Lung Cancers by Arm

	LDCT Arm	CXR Arm	
	# (%)	# (%)	
All	1701	1681	
Histology			**P-value** [3]
All NSCLC	1397 (82.1)	1343 (79.9)	0.28
BAC	121 (7.1)	46 (2.7)	<0.0001
Adenocarcinoma	608 (35.7)	598 (35.6)	0.76
Squamous	416 (24.5)	395 (23.5)	0.45
Large Cell	56 (3.3)	53 (3.2)	0.77
Other NSCLC	196 (11.5)	251 (14.9)	0.009
SCLC	245 (14.4)	291 (17.3)	0.05
Carcinoid	12 (0.7)	7 (0.4)	
Unknown	47 (2.8)	40 (2.4)	
Stage [1]			
I [2]	673 (39.6)	462 (27.5)	<0.0001
IA	523	326	
1B	148	134	
II [2]	145 (8.5)	153 (9.1)	0.65
IIA	91	80	
IIB	43	66	
III [2]	298 (17.5)	321 (19.1)	0.36
IIIA	204	216	
IIIB	84	94	
IV	468 (27.5)	597 (35.5)	<0.0001
Occult	5	4	
Unknown	112 (6.6)	143 (8.5)	

Note: ICD-O-3 8000 considered unknown; 8010 considered Other NSCLC.

[1]. Based on AJCC 6th edition for cases through 2009 and (primarily) 7th edition for cases from 2010 on.

[2]. Includes some cases without A,B distinction.

[3]. For difference in proportion of cases.

Table 3.

Lung cancer mortality, stage IV incidence and overall mortality by arm

	LDCT	CXR	Difference across arms (95% CI) [CXR minus LDCT]	RR (95% CI)	P-value Interaction [2]
All lung cancer deaths	# (per 1,000 subjects)	# (per 1,000 subjects)	per 1,000 subjects		
All subjects	1147 (42.9)	1236 (46.2)	3.3 (−0.2 – 6.8)	0.92 (0.85-1.00)	
Men	733 (46.5)	755 (47.9)	1.4 (−3.3 – 6.1)	0.97 (0.87-1.07)	0.17
Women	414 (37.8)	481 (43.9)	6.1 (0.8 – 11.3)	0.86 (0.75-0.98)	
Current Smoker	724 (56.3)	818 (63.4)	7.1 (1.3-12.9)	0.88 (0.80-0.97)	0.12
Former Smoker	423 (30.5)	418 (30.2)	−0.3 (−4.3-3.8)	1.01 (0.88-1.15)	
Age 55-64 at randomization	641 (32.7)	739 (37.7)	5.0 (1.3-8.6)	0.86 (0.78-0.96)	0.051
Age 65-74 at randomization	506 (71.2)	497 (69.9)	−1.3 (−9.7-7.2)	1.01 (0.90-1.15)	
Lung cancer deaths - dilution-adjusted analysis [1]					
All subjects	578 (21.6)	646 (24.2)	2.5 (0.001-5.1)	0.89 (0.80-0.997)	
Men	373 (23.7)	390 (24.7)	1.1 (−2.3 – 4.5)	0.95 (0.83-1.10)	0.14
Women	205 (18.7)	256 (23.3)	4.6 (0.8 – 8.4)	0.80 (0.66-0.96)	
Current Smoker	356 (27.7)	423 (32.8)	5.1 (0.9-9.3)	0.84 (0.73 −0.97)	0.16
Former Smoker	222 (16.0)	223 (16.1)	0.1 (−2.9-3.1)	0.99 (0.82-1.19)	
Age 55-64 at randomization	310 (15.8)	362 (18.4)	2.6 (0.1-5.2)	0.85 (0.73-0.99)	0.39
Age 65-74 at randomization	268 (37.7)	284 (40.0)	2.3 (−4.1-8.6)	0.94 (0.80-1.11)	
Stage IV cases					
All subjects	468 (17.5)	597 (22.3)	4.8 (2.5 – 7.2)	0.79 (0.70-0.89)	
Men	303 (19.2)	365 (23.2)	3.9 (0.8 – 7.1)	0.83 (0.71-0.97)	0.24
Women	165 (15.1)	232 (21.2)	6.1 (2.6 – 9.6)	0.71 (0.58-0.87)	
Current Smoker	297 (23.1)	386 (29.9)	6.8 (2.9-10.7)	0.77 (0.66-0.90)	0.69
Former Smoker	171 (12.3)	211 (15.3)	2.9 (0.2-5.7)	0.81 (0.66-0.99)	
Age 55-64 at randomization	278 (14.2)	367 (18.7)	4.5 (2.0-7.0)	0.76 (0.65-0.89)	0.48
Age 65-74 at randomization	190 (26.7)	230 (32.4)	5.6 (0.1-11.2)	0.83 (0.69-1.01)	
Stage IV cases through year 6					
All Subjects	245 (9.2)	344 (12.9)	3.7 (1.9 – 5.5)	0.71 (0.60-0.84)	
Men	165 (10.5)	214 (13.6)	3.1 (0.7-5.5)	0.77 (0.63-0.95)	0.21
Women	80 (7.3)	130 (11.9)	4.5 (2.0 – 7.1)	0.62 (0.47-0.82)	
Current Smoker	153 (11.9)	221 (17.1)	5.2 (2.3-8.2)	0.70 (0.57-0.86)	0.66
Former Smoker	92 (6.6)	123 (8.9)	2.3 (0.2-4.3)	0.75 (0.57-0.98)	
Age 55-64	140 (7.1)	207 (10.5)	3.4 (1.6-5.3)	0.68 (0.55-0.84)	0.46
Age 65-74	105 (14.8)	137 (19.3)	4.5 (0.3-8.8)	0.77 (0.60-0.99)	
Overall mortality (all subjects)	5253 (196.6)	5366 (200.7)	4.2 (−2.6-10.9)	0.97 (0.94-1.01)	

	LDCT	CXR	Difference across arms (95% CI) [CXR minus LDCT]	RR (95% CI)	P-value Interaction [2]
Overall mortality excluding lung cancer deaths (all subjects)	4106 (153.7)	4130 (154.5)	0.9 (−5.3-7.0)	0.99 (0.95-1.03)	

[1] All deaths with corresponding lung cancer diagnosis within 6 years of randomization were included.

[2] P-value for interaction of trial arm by age, sex, or smoking status for the RR.

Table 4.

Lung cancer mortality results for comparable median follow-up periods based on calendar-time versus study-time cutoffs.

Time Period for inclusion of lung cancer deaths	Median (25th/75th) years follow-up for mortality	LDCT	CXR	RR (95% CI)	Difference per 1,000 subjects
Study Time Cutoff		Lung Cancer Deaths	Lung Cancer Deaths		
Through Study Year					
5.0	5.0 (5.0/5.0)	312	370	0.84 (0.72-0.98)	2.2
5.5	5.5 (5.5/5.5)	347	427	0.81 (0.71-0.93)	3.0
6.0	6.0 (6.0/6.0)	398	491	0.81 (0.71-0.92)	3.5
6.5	6.5 (6.5/6.5)	457	550	0.83 (0.73-0.94)	3.5
7.0	7.0 (7.0/7.0)	517	600	0.86 (0.76-0.97)	3.1
Calendar Time Cutoff					
Through Jan 15, 2009	5.5 (5.2/5.9)	356	443	0.80 (0.73-0.93)	3.3
Through Dec 31, 2009	6.5 (6.1/6.9)	469	552	0.84 (0.75-0.96)	3.1

Note: Includes only those lung cancer deaths occurring in the given time periods. This is in contrast to the dilution analyses, where deaths can occur any time but diagnoses have to occur during certain time periods.