Published in final edited form as:

Int J Cancer. 2022 November 01; 151(9): 1491–1501. doi:10.1002/ijc.34167.

The influence of post-screening follow-up time and participant characteristics on estimates of overdiagnosis from lung cancer screening trials

Mengmeng Li, MD, PhD[1,†], **Li Zhang, MD, PhD**[2,†], **Hadrien Charvat, MD, PhD**[2,3,4], **Matthew E Callister, PhD**[5], **Peter Sasieni, PhD**[6], **Evangelia Christodoulou, PhD**[7,8], **Rudolf Kaaks, PhD**[7,8], **Mattias Johansson, PhD**[2], **Andre L. Carvalho, MD, PhD**[2], **Salvatore Vaccarella, PhD**[2], **Hilary A. Robbins, PhD**[2]

[1]Department of Cancer Prevention, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou, China

[2]International Agency for Research on Cancer, Lyon, France

[3]Faculty of International Liberal Arts, Juntendo University, Tokyo, Japan

[4]Division of International Health Policy Research, Institute for Cancer Control, National Cancer Center, Tokyo, Japan

[5]Leeds Teaching Hospitals, Leeds, United Kingdom

[6]Kings College London, London, United Kingdom

[7]Division of Cancer Epidemiology, German Cancer Research Center, Heidelberg, Germany

[8]Translational Lung Research Center (TLRC) Heidelberg, Member of the German Center for Lung Research (DZL), Heidelberg, Germany

Correspondence: Dr. Mengmeng Li, Department of Cancer Prevention, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou, China. Tel: +86 020-87345679; limm@sysucc.org.cn. Dr. Hilary A. Robbins, Genomic Epidemiology Branch, International Agency for Research on Cancer, 150 cours Albert Thomas, Lyon, France. Tel: +33 72 73 80 88; robbinsh@iarc.fr. Dr. Hadrien Charvat, Division of International Health Policy Research, Institute for Cancer Control, National Cancer Center, Tokyo, Japan. Tel: +81 3-3547-5201; chadrien@ncc.go.jp.

[†]Mengmeng Li and Li Zhang contributed equally to this work.

Author Contributions: Mengmeng Li: Conceptualization, Data curation, Investigation, Methodology, Formal analysis, Software, Data interpretation, Writing - original draft, review & editing. **Li Zhang:** Data curation, Investigation, Data interpretation, Writing - original draft, review & editing. **Hadrien Charvat:** Investigation, Methodology, Statistic analysis, Software, Data interpretation, Writing - review & editing. **Matthew E Callister:** Investigation, Data interpretation, Writing - review & editing. **Peter Sasieni:** Investigation, Data interpretation, Writing - review & editing. **Evangelia Christodoulou:** Data curation, Data interpretation, Writing - review & editing. **Rudolf Kaaks:** Data curation, Data interpretation, Writing - review & editing. **Mattias Johansson:** Investigation, Data interpretation, Writing - review & editing. **Andre L. Carvalho:** Investigation, Data interpretation, Writing - review & editing. **Salvatore Vaccarella:** Investigation, Data interpretation, Writing - review & editing. **Hilary A. Robbins:** Conceptualization, Data curation, Investigation, Methodology, Data interpretation, Writing - original draft, review & editing. The work reported in the paper has been performed by the authors, unless clearly specified in the text.

Conflicts of Interest: None declared.

Disclaimer: Where authors are identified as personnel of the International Agency for Research on Cancer / World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer / World Health Organization.

Ethics Statement: Our study involved only secondary analysis of de-identified data and is classified as exempt from IRB approval by the NIH Common Rule.

Abstract

We aimed to explore the underlying reasons that estimates of overdiagnosis vary across and within low-dose CT (LDCT) lung cancer screening trials. We conducted a systematic review to identify estimates of overdiagnosis from randomised controlled trials of LDCT screening. We then analysed the association of Ps (the excess incidence of lung cancer as a proportion of screen-detected cases) with post-screening follow-up time using a linear random effects meta-regression model. Separately, we analysed annual Ps estimates from the US National Lung Screening Trial (NLST) and German Lung Cancer Screening Intervention Trial (LUSI) using exponential decay models with asymptotes. We conducted stratified analyses to investigate participant characteristics associated with Ps using the extended follow-up data from NLST. Among 12 overdiagnosis estimates from 8 trials, the post-screening follow-up ranged from 3.8 to 9.3 years, and Ps ranged from −27.0% (ITALUNG, 8.3y follow-up) to 67.2% (DLCST, 5.0y follow-up). Across trials, 39.1% of the variation in Ps was explained by post-screening follow-up time. The annual changes in Ps were −3.5% and −3.9% in the NLST and LUSI trials, respectively. Ps was predicted to plateau at 2.2% for NLST and 9.2% for LUSI with hypothetical infinite follow-up. In NLST, Ps increased with age from −14.9% (55–59 years) to 21.7% (70–74 years), and time trends in Ps varied by histological type. The findings suggest that differences in post-screening follow-up time partially explain variation in overdiagnosis estimates across lung cancer screening trials. Estimates of overdiagnosis should be interpreted in the context of post-screening follow-up and population characteristics.

Keywords

Overdiagnosis; lung cancer screening; randomised controlled trial

Introduction

Overdiagnosis in cancer screening refers to the diagnosis of a cancer that would otherwise not cause clinical symptoms or death[1], which causes unneeded treatment, additional cost, and physical and mental harms. There are two primary underlying factors leading to overdiagnosis: indolent cancers that are less likely to progress and competing causes of death that occur before cancer symptoms arise[1–3]. The amount of overdiagnosis is commonly estimated using data from randomised controlled trials (RCTs) by comparing the number of cancers diagnosed in the screening arm with the number diagnosed in the control arm.

Overdiagnosis of lung cancer during low-dose computed tomography (LDCT) screening remains controversial, largely because estimates of its magnitude have varied widely across RCTs, ranging from no evidence of overdiagnosis[4] to 67% of screen-detected lung cancers being considered overdiagnosed[5]. This variability has led to extensive debate regarding the net harms, and therefore the balance between benefits and harms, of LDCT lung cancer screening. Long-term follow-up of randomised trials offers the opportunity to estimate overdiagnosis as the amount of excess incidence among the screen-detected lung cancers. With sufficient follow-up time to allow the cases in the non-screened arm to “catch up” to the cancer cases detected earlier in the screened arm[1], the excess incidence eventually

approximates the amount of true overdiagnosis. When the post-screening follow-up is short, overdiagnosis cannot be distinguished from the lead time effect[6, 7].

We hypothesised that a portion of the heterogeneity in overdiagnosis estimates between RCTs of lung cancer screening could be attributable to the fact that they were calculated after different amounts of post-screening follow-up time. Moreover, the role of follow-up time and participant characteristics in explaining overdiagnosis estimates within individual trials remains largely unexplored. A detailed description of these influences could improve our understanding of the true magnitude of overdiagnosis in lung cancer screening and identify the means to reduce it. In this study, we aimed to identify factors that may explain the variation in overdiagnosis estimates between and within RCTs of lung cancer screening. We reviewed data from published randomised trials of LDCT screening and conducted a secondary analysis of the National Lung Screening Trial (NLST) and German Lung Cancer Screening Intervention Trial (LUSI) with extended follow-up data.

Methods

Overview

There are two indicators[8, 9] commonly used to represent overdiagnosis in RCTs with a stop-screening design[6]: Ps, the ratio between excess incidence (calculated as the difference between the cumulative incidence of lung cancer in the screening and control arms) and the cumulative incidence of screen-detected lung cancers; and Pa, the ratio between excess incidence and the cumulative incidence of all lung cancers in the screening arm. Conceptually, Ps is a better representation of overdiagnosis than Pa, because if we refer to someone as being overdiagnosed by screening, the cancer must be screen-detected. Therefore, in this study we considered both Ps and Pa, but used Ps as the primary outcome measure.

Review of lung cancer screening trials and analysis of association between follow-up time and overdiagnosis estimates across trials

We systematically searched three databases (Embase, Pubmed, and Web of Science) for studies reporting RCTs of LDCT lung cancer screening. The detailed search strategy, selection criteria, and the flow chart of the selection process are shown in Supplementary Box S1 and Figure S1. A total of 13 articles describing 12 overdiagnosis estimates were included. For each of the selected studies, we extracted the following parameters: the number of participants in the screening arm (T_S) and control arm (T_C); the number of lung cancer cases accrued at the end of follow-up in the screening arm (D_S) and control arm (D_C); and screen-detected lung cancer cases in the screening arm (SD_S). Ps and Pa were then calculated as[10]:

$$P_S = \frac{\left(\frac{D_S}{T_S} - \frac{D_C}{T_C}\right)}{\frac{SD_S}{T_S}}$$

$$P_a = \frac{\left(\frac{D_S}{T_S} - \frac{D_C}{T_C}\right)}{\frac{D_S}{T_S}}$$

The variances of Ps and Pa were estimated by Monte-Carlo simulations (see Supplementary Table S1). For studies where post-screening follow-up time was not given, it was estimated as the difference between reported median follow-up years and the screening years, or between reported follow-up years and median duration of screening, whichever was available. The impact of post-screening follow-up time on the values of Ps and Pa was assessed using a linear random effects meta-regression model accounting for the multilevel structure of the data (i.e., several studies contributed two data points) and fitted by maximum likelihood. We calculated the proportion of variation in Ps and Pa estimates that was explained by post-screening follow-up time as the proportional reduction in residual heterogeneity when using post-screening follow-up time as a moderator in the meta-regression model, compared to fitting the model without a moderator[11].

Analyses of the association between Ps and Pa estimates and follow-up time within the NLST and LUSI trials

For the NLST and LUSI trials, we had access to detailed data describing the cumulative incidence of lung cancer in the screening and control arms, together with the cumulative incidence of screen-detected cancers, at annual timepoints following the end of screening. Detailed descriptions of the two trials have been published previously[12, 13]. Briefly, the NLST recruited former or current smokers aged 55–74 years and randomised the participants to three annual rounds of either LDCT or chest X-ray screening, and the LUSI trial randomised 50–69-year-olds with a history of heavy smoking to five annual rounds of LDCT screening or usual care. At the time of this study, follow-up data was available until a median of 9.3 and 6.7 years since the last screen for the NLST and LUSI trials, respectively[9, 12]. We then calculated the Ps and Pa estimates at each annual timepoint and used linear regression models to assess the relationship between Ps/Pa estimates and follow-up time for each trial. The difference between the time trend across trials was examined using a Z-test with the formula $z = \frac{\beta_1 - \beta_2}{\sqrt{SE_1^2 + SE_2^2}}$, where β and SE represent the slopes and standard errors in the two trials[14]. For the NLST trial, we also conducted stratified analyses by histological subtype.

For Ps, we also fit exponential decay models with asymptotes to both studies combined, which assume that Ps eventually reaches a plateau after long-term follow-up. We made this assumption because the excess incidence between the screening and control arms should theoretically stabilise after sufficiently long follow-up time, and that is when overdiagnosis can be validly measured[6]. This model includes parameters for the initial Ps level, the rate of exponential decrease, and the plateau Ps value. We assumed the rate of decrease of Ps was similar between the studies, because it reflects characteristics of the disease, in particular the maximum preclinical period[6], and because the slopes in the linear regression models

were similar between the two studies. Therefore, we used a common exponential decrease parameter while the other parameters were study-specific.

Analysis of participant characteristics associated with Ps and Pa estimates in the extended follow-up of NLST

We used the extended follow-up data of NLST with a median duration of 9.3 years since the last screen[12], the longest duration among trials, to calculate Ps and Pa after stratifying by various characteristics including sex, age (in 5-year groups), race (white or non-white), smoking status (former or current smoker), chronic obstructive pulmonary disease (COPD), predicted lung cancer risk at baseline, comorbid conditions, life expectancy, and histological subtype. Lung cancer risk was calculated by the $PLCO_{M2012}$ model[15], which predicts the probability of lung cancer diagnosis during a 6-year period. Comorbidity for each participant was measured by the Charlson comorbidity index[16, 17], adapted to the availability of information in the NLST, which combines the number and the severity of comorbid disease into a weighted index. Life expectancy without lung cancer screening was estimated using the Life Years Gained From Screening-CT (LYFS-CT) model[18]. Confidence intervals were obtained by taking the 2.5th and 97.5th percentiles of the empirical Ps/Pa distribution generated for each stratum according to the Monte-Carlo procedure mentioned above. Statistical tests for the difference or trend in Ps and Pa estimates for each variable were conducted using a simulation-based approach for heterogeneity test (P_{Hetero}) and weighted linear regression models for trend test (P_{Trend}) (Supplementary Box S2).

All analyses were conducted in R (version 3.6.3). The “metafor” R package[19] was used to fit multilevel random effects meta-analysis and meta-regression models. Calculations for the LYFS-CT model were performed using the “lcmodels” R package[18].

Results

Review of lung cancer screening trials and analysis of the association between Ps and Pa estimates and follow-up time across different trials

We identified 8 lung cancer screening trials with LDCT that met our inclusion criteria: NLST[12, 20, 21], Nederlands–Leuvens Longkanker Screenings Onderzoek (NELSON)[22], Danish Lung Cancer Screening Trial (DLCST)[5, 23], Multicentric Italian Lung Detection (MILD)[24], LUSI[9, 13], UK Lung Cancer Screening Trial (UKLS)[25], Italian Lung Cancer Screening Trial (ITALUNG)[4, 26], and Detection And screening of early lung cancer with Novel imaging Technology (DANTE)[27] (Table 1). For 4 trials, Ps and Pa estimates were available at two timepoints. Follow-up time after the final screen ranged from 3.8 years (MILD) to 9.3 years (NLST extended follow-up). Ps estimates varied from −27.0% (ITALUNG, 8.3y follow-up) to 67.2% (DLCST, 5.0y follow-up), and Pa estimates ranged from −11.3% to 44.8% also between ITALUNG and DLCST. The parameters used in the regression model between Ps and Pa estimates and follow-up years after the final screen are shown in Supplementary Table S1. Across different trials, estimates of Ps decreased with longer post-screening follow-up time, with an annual change of −4.0% (95%CI: −8.1%, 0.1%) (Figure 1A). The post-screening follow-up time was estimated to explain 39.1% of the variation in Ps estimates across trials. For Pa, estimates declined by −2.4% (95%CI:

−4.4%, −0.5%) annually and post-screening follow-up time accounted for 42.3% of the variation (Figure 1B).

Analysis of the association between Ps and Pa estimates and follow-up time within the NLST and LUSI trials

Within both the NLST and LUSI trials, when applying linear regression models, Ps estimates declined continuously with follow-up after the last screen (Figure 2). For NLST, the value of Ps decreased by an absolute amount of −3.5% annually (95%CI: −4.0%, −3.0%) from 1.5 to 9.3 years post-screening follow-up, and the R-squared was 95.6%. For LUSI, the annual absolute change in Ps estimates was similar to NLST at −3.9% (95%CI: −5.7%, −2.0%) from 0.9 to 6.7 years post-screening follow-up, with an R-squared of 71.5%. There was no statistically significant difference in the slopes between the two trials (Ps: $P = 0.734$; Pa: $P = 0.295$). At each follow-up timepoint from 2 to 6 years after the last screen, the Ps estimates in NLST were approximately 3–4% lower than in LUSI; for example, the regression prediction at 4 years of follow-up was 22.1% in NLST compared with 25.8% in LUSI (Figure 2). When fitting data from the two trials with exponential decay models with an asymptote, we estimated that Ps would plateau at approximately 2.2% for NLST and 9.2% for LUSI with hypothetical infinite follow-up time (Figure 3). For Pa, the annual absolute change was −3.2% (95%CI: −4.1%, −2.4%) and −4.2% (95%CI: −5.9%, −2.6%) for NLST and LUSI, respectively (Supplementary Figure S2).

Figure 4 shows histology-specific Ps estimates versus follow-up time since the final screen in NLST. The time-trend in Ps varied across histological types ($P < 0.0001$, Supplementary Table S2). Excess incidence of all adenocarcinomas declined over time (annual change of −2.8%, 95%CI: −3.6%, −2.1%). Considering bronchioloalveolar carcinoma (BAC) and non-BAC adenocarcinomas separately, Ps for BAC was stable over time whereas Ps for non-BAC adenocarcinomas declined steadily, with an annual change of −4.0% (95%CI: −4.9%, −3.1%). For squamous cell carcinoma, Ps was lower than for adenocarcinoma, but was approximately stable over time with an annual change of −0.8% (95%CI: −2.3%, 0.6%). The excess incidence of small cell carcinoma showed a strong decline over time, with an annual change of −13.0% (95%CI: −15.9%, −10.1%). For large cell carcinoma, Ps remained stably low over time. The directions of the time trends in Pa by subtype were similar to Ps (Supplementary Figure S3), but the slopes were smaller (Supplementary Table S2).

Analysis of participant characteristics associated with Ps and Pa estimates in the extended follow-up of NLST

Table 2 shows Ps and Pa estimates from NLST after 9.3 years of median follow-up after the last screen, along with absolute excess incidence rates. Overall, the excess incidence rate of lung cancer was 0.8 per 1,000 people, Ps was 3.3% (95%CI: −14.2%, 19.8%), and Pa was 1.3% (95%CI: −5.5%, 7.5%). By age, the excess incidence rate increased steadily from −2.3 cases per 1,000 (age 55–59 years) to 0.6, 3.9, and 10.2 (70–74 years) across 5-year age groups, with Ps ranging from −14.9% (55–59 years) to 2.5, 10.9, and 21.7% (70–74 years) ($P = 0.015$). A similar trend was observed for competing (non-lung-cancer) mortality, which ranged from 9.0 to 30.2 deaths per 100 people across age groups (Supplementary Figure S4). Although not statistically significant, higher values of Ps were observed among people with

shorter life expectancy (19.4% for ≤10 years life expectancy versus −13.8% for >20 years, $P = 0.13$). By histological subtype, the Ps estimate for BAC was 79.0% (95%CI: 59.6%, 96.4%), and a small value was found for adenocarcinoma excluding BAC (4.4%; 95%CI: −23.1%, 29.1%). The results for Pa were similar to those for Ps (Table 2).

Discussion

Our study was motivated by the wide variation in published estimates of overdiagnosis in lung cancer screening and the resulting debate among clinicians and researchers. We sought to assess whether this variation could be explained by differences in the length of post-screening follow-up and the impact of participant characteristics on the estimates within trials. We found that 39% and 42% of the variation in Ps and Pa estimates across trials were explained by post-screening follow-up, and that Ps and Pa estimates decreased by 3–4% annually with post-screening follow-up time in both the NLST and LUSI trials. Stratified analysis found that excess incidence increased with age at screening initiation and that time-trends in excess incidence varied by histological type.

Our analysis demonstrated that the length of follow-up after the final screen may explain approximately 39% of the variation in excess lung cancer incidence across LDCT screening trials. Two trials fell outside the pattern of the others, including the DLCST, which gave an unusually high Ps estimate of 67%, and the ITALUNG trial, which produced negative Ps estimates. We speculate that the DLCST result may be caused by higher-risk participants being randomly assigned more frequently to the screening arm[28], and the negative estimates in ITALUNG might be partly due to the higher rate of quitting smoking in the screening arm compared with the control arm[29]. There are different factors that might explain the remaining 61% of the variation in overdiagnosis estimates, including differences across trials in the number of screens and screening intervals, screening sensitivity, nodule management protocols, participant characteristics, and random chance. Unfortunately, due to the small number of trials, lack of wide variability across trials, and heterogeneous reporting of information across trials, we found it was not possible to directly assess or adjust for the influence of design factors, nor of trial-level differences in participant characteristics (e.g. average age).

We found that the annual decline in both Ps and Pa estimates during post-screening follow-up occurred at a similar rate in the NLST and LUSI trials, which was consistent with the annual decline observed across all trials (3–4%). However, the estimates in LUSI were higher than in NLST at the same follow-up timepoints. This could be because NLST had chest radiography screening in the control group, which may have increased lung cancer incidence and led to a lower apparent amount of excess incidence due to LDCT screening, whereas LUSI had no screening in the control arm. Another possible explanation is the higher number of screening rounds in LUSI compared with NLST[13, 20]. Under hypothetical infinite follow-up, as shown by our exponential decay model, the excess incidence was predicted to stabilise at around 2% for NLST and 9% for LUSI, which may estimate the true amount of overdiagnosis in these two trials.

Our results on overdiagnosis by histological type are generally consistent with those found in microsimulation studies, with BAC ranking the highest and small cell carcinoma the lowest in amount of overdiagnosis[30]. The high and stable proportion of overdiagnosis in BAC may be related to the indolent nature of most BACs as compared to other subtypes, or that the corresponding lesions in non-screened participants may present as invasive adenocarcinoma after sufficient follow-up time. We note that the histological term BAC was superseded in 2011 by the categories of adenocarcinoma in situ (AIS), minimally invasive adenocarcinoma (MIA), and lepidic predominant adenocarcinoma[31]. AIS and MIA are recognized as precursors to fully invasive adenocarcinomas and commonly equate to pure ground glass and part solid nodules on CT scans. Despite these changes to the classification of lung adenocarcinoma, the same morphology codes were used to define BAC throughout the NLST follow-up period[12]. The reason for the steep decline over time in the Ps estimate for small cell carcinoma is unclear, but its large negative values suggest that the likelihood for a small cell carcinoma to be overdiagnosed is much smaller compared with adenocarcinoma, squamous cell carcinoma and large cell carcinoma.

Our analysis suggests that overdiagnosis in lung screening is likely to be higher among older people and those with shorter life expectancy, which is expected due to the influence of competing mortality. Although these patterns were apparent within a single trial, it is difficult to assess whether they are important reasons for heterogeneity in overdiagnosis estimates across trials, because the mean age of participants is similar across many trials and individual-level data are not readily available. For older individuals or people with shorter life expectancy, a higher likelihood of overdiagnosis may be one element to incorporate into shared decision-making conversations[32]. Another strategy is to incorporate comorbidities into screening decisions[33]. Ultimately, a holistic approach incorporating individual life expectancy, comorbidities, and personal preferences[34] is likely needed for people for whom potential screening benefits do not clearly outweigh potential harms.

One limitation of our study is the heterogeneity in study designs across different lung screening trials, including differences in the number and interval of screens and whether there was an intervention in the control arm. Since we relied on data from randomised trials, our analysis cannot fully represent ‘real world’ screening where screening participants tend to be less healthy than participants in trials[35]. We note that small inconsistencies may exist in the calculation of the length of post-screening follow-up between trials, as some were estimated using the median follow-up among participants and others using the study year. Confidence intervals were wide for some estimates despite using data from the largest available trial (NLST), since overdiagnosis can only be measured at a group level and not for individuals.

Neither the NLST nor the LUSI results indicate that a clear plateau in the Ps estimate has been reached. In the Mayo Lung Project trial of chest radiography and sputum cytology screening, after approximately 16 years of post-screening follow-up, excess lung cancers still persisted in the screening arm[36]. For breast cancer, in the Canadian National Breast Screening Study, the Ps estimate decreased from 29% at the end of the screening period to 22% after ten years and remained constant over the following two decades[37]. Among the RCTs evaluated by the Independent UK Panel on Breast Cancer Screening, the Ps estimates

ranged from 20% to 29%[38]. Based on a screening cohort, a recent study found 15% of screen-detected breast cancers in the US were overdiagnosed[39]. Considering the NLST and LUSI plateau estimates from exponential decay models (2.2% and 9.2%, respectively), and similarly low proportions of lung cancer overdiagnosis in microsimulation studies – where the mean Ps estimate ranged from 5.6–6.3% across 4 models within the Cancer Intervention and Surveillance Modeling Network[40] – this comparison suggests that overdiagnosis in lung cancer screening may be small relative to breast cancer screening. However, the wide variation in overdiagnosis estimates for breast cancer screening across studies[41] makes it difficult to reach a definite conclusion.

Our findings provide a new perspective on the heterogeneous estimates of overdiagnosis in lung cancer screening. We found that estimates from most trials are not necessarily inconsistent with one another, provided they are interpreted in the context of how much follow-up time had elapsed before they were calculated. Therefore, if one aims to estimate overdiagnosis from a trial with fewer than 10 years of follow-up, it is important to look at excess diagnoses as a function of time since the final screen, to discuss the trend, and to extrapolate what the excess might be thereafter. Analyses within the NLST and LUSI trials confirmed the strong influence of follow-up time, showed that overdiagnosis is more likely among individuals of older age, and illustrated heterogeneous patterns among histological subtypes. Taken together, our findings emphasise that the probability of overdiagnosis does not take on a single value and that estimates of overdiagnosis must be interpreted in the context of study design and participant characteristics.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

We thank Gordon Jones, Brian Barrett, and Jerome Mabie at Information Management Services, Inc. for providing data from the National Lung Screening Trial.

Funding:

This project had no specific funding. Mengmeng Li was supported by Guangdong Basic and Applied Basic Research Foundation (2021A1515110625). Hilary A. Robbins and Mattias Johansson were supported by the INTEGRAL project (NCI U19 CA203654). H. Robbins was additionally supported by NCI R03 CA245979.

Data Availability Statement:

Much of the data analysed in our study was extracted from existing publications. The National Lung Screening Trial data are available to researchers upon request to the US National Cancer Institute. The LUSI Trial data are not publicly available, but interested researchers may contact the German Cancer Research Center (DKFZ). Further information is available from the corresponding author upon request.

List of abbreviations

AIS	adenocarcinoma in situ

BAC	bronchioloalveolar carcinoma
COPD	chronic obstructive pulmonary disease
CXR	chest X-ray
DANTE	Detection And screening of early lung cancer with Novel imaging TEchnology
DLCST	Danish Lung Cancer Screening Trial
ITALUNG	Italian Lung Cancer Screening Trial
LDCT	low-dose computed tomography
LUSI	German Lung Cancer Screening Intervention Trial
LYFS-CT	Life Years Gained From Screening-CT
MIA	minimally invasive adenocarcinoma
MILD	Multicentric Italian Lung Detection
NELSON	Nederlands–Leuvens Longkanker Screenings Onderzoek
NLST	National Lung Screening Trial
RCT	randomised controlled trial
UKLS	UK lung cancer screening trial

References

1. Welch HG, Black WC. Overdiagnosis in cancer. J Natl Cancer Inst 2010;102:605–13. [PubMed: 20413742]
2. Callister MEJ, Sasieni P, Robbins HA. Overdiagnosis in lung cancer screening. Lancet Respir Med 2021;9:7–9. [PubMed: 33412118]
3. Carter JL, Coletti RJ, Harris RP. Quantifying and monitoring overdiagnosis in cancer screening: a systematic review of methods. BMJ 2015;350:g7773. [PubMed: 25569206]
4. Paci E, Puliti D, Lopes Pegna A, Carrozzi L, Picozzi G, Falaschi F, Pistelli F, Aquilini F, Ocello C, Zappa M, Carozzi FM, Mascalchi M, et al. Mortality, survival and incidence rates in the ITALUNG randomised lung cancer screening trial. Thorax 2017;72:825–31. [PubMed: 28377492]
5. Heleno B, 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 Intern Med 2018;178:1420–22. [PubMed: 30276415]
6. Gulati R, Feuer EJ, Etzioni R. Conditions for valid empirical estimates of cancer overdiagnosis in randomized trials and population studies. Am J Epidemiol 2016;184:140–7. [PubMed: 27358266]
7. Ten Haaf K, de Koning HJ. Overdiagnosis in lung cancer screening: why modelling is essential. J Epidemiol Community Health 2015;69:1035–9. [PubMed: 26071497]
8. Patz EF Jr., Pinsky P, Gatsonis C, Sicks JD, Kramer BS, Tammemägi MC, Chiles C, Black WC, Aberle DR. Overdiagnosis in low-dose computed tomography screening for lung cancer. JAMA Intern Med 2014;174:269–74. [PubMed: 24322569]
9. Gonzalez Maldonado S, Motsch E, Trotter A, Kauczor HU, Heussel CP, Hermann S, Zeissig SR, Delorme S, Kaaks R. Overdiagnosis in lung cancer screening: Estimates from the German Lung Cancer Screening Intervention Trial. Int J Cancer 2021; 148(5):1097–105. [PubMed: 32930386]

10. Brodersen J, Voss T, Martiny F, Siersma V, Barratt A, Heleno B. Overdiagnosis of lung cancer with low-dose computed tomography screening: meta-analysis of the randomised clinical trials. Breathe 2020;16:200013. [PubMed: 32194774]
11. Xu R Measuring explained variation in linear mixed effects models. Stat Med 2003;22:3527–41. [PubMed: 14601017]
12. National Lung Screening Trial Research Team. Lung cancer incidence and mortality with extended follow-up in the National Lung Screening Trial. J Thorac Oncol 2019;14:1732–42. [PubMed: 31260833]
13. Becker N, Motsch E, Trotter A, Heussel CP, Dienemann H, Schnabel PA, Kauczor HU, Maldonado SG, Miller AB, Kaaks R, Delorme S. Lung cancer mortality reduction by LDCT screening-Results from the randomized German LUSI trial. Int J Cancer 2020;146:1503–13. [PubMed: 31162856]
14. Yang J, Zhou M, Li M, Yin P, Hu J, Zhang C, Wang H, Liu Q, Wang B. Fine particulate matter constituents and cause-specific mortality in China: A nationwide modelling study. Environ Int 2020;143:105927. [PubMed: 32619910]
15. Tammemagi MC, Katki HA, Hocking WG, Church TR, Caporaso N, Kvale PA, Chaturvedi AK, Silvestri GA, Riley TL, Commins J, Berg CD. Selection criteria for lung-cancer screening. N Engl J Med 2013;368:728–36. [PubMed: 23425165]
16. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373–83. [PubMed: 3558716]
17. Kastner C, Armitage J, Kimble A, Rawal J, Carter PG, Venn S. The Charlson comorbidity score: a superior comorbidity assessment tool for the prostate cancer multidisciplinary meeting. Prostate Cancer Prostatic Dis 2006;9:270–4. [PubMed: 16770340]
18. Cheung LC, Berg CD, Castle PE, Katki HA, Chaturvedi AK. Life-gained-based versus risk-based selection of smokers for lung cancer screening. Ann Intern Med 2019;171:623–32. [PubMed: 31634914]
19. Viechtbauer W Conducting meta-analyses in R with the metafor package. J Stat Softw 2010;36:1–48.
20. National Lung Screening Trial Research T, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, Gareen IF, Gatsonis C, Marcus PM, Sicks JD. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 2011;365:395–409. [PubMed: 21714641]
21. Patz EF, Pinsky P, Gatsonis C, Sicks JD, Kramer BS, Tammemägi MC, Chiles C, Black WC, Aberle DR. Overdiagnosis in low-dose computed tomography screening for lung cancer. JAMA Intern Med 2014;174:269–74. [PubMed: 24322569]
22. de Koning HJ, van der Aalst CM, de Jong PA, Scholten ET, Nackaerts K, Heuvelmans MA, Lammers JJ, Weenink C, Yousaf-Khan U, Horeweg N, van 't Westeinde S, Prokop M, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. N Engl J Med 2020;382:503–13. [PubMed: 31995683]
23. Wille MM, Dirksen A, Ashraf H, Saghir Z, Bach KS, Brodersen J, Clementsen PF, Hansen H, Larsen KR, Mortensen J, Rasmussen JF, Seersholm N, et al. Results of the randomized Danish lung cancer screening trial with focus on high-risk profiling. Am J Respir Crit Care Med 2016;193:542–51. [PubMed: 26485620]
24. Pastorino U, Silva M, Sestini S, Sabia F, Boeri M, Cantarutti A, Sverzellati N, Sozzi G, Corrao G, Marchianò A. Prolonged lung cancer screening reduced 10-year mortality in the MILD trial: new confirmation of lung cancer screening efficacy. Ann Oncol 2019;30:1162–9. [PubMed: 30937431]
25. Field JK, Vulkan D, Davies MPA, Baldwin DR, Brain KE, Devaraj A, Eisen T, Gosney J, Green BA, Holemans JA, Kavanagh T, Kerr KM, et al. Lung cancer mortality reduction by LDCT screening: UKLS randomised trial results and international meta-analysis. Lancet Reg Health Eur 2021;10:100179. [PubMed: 34806061]
26. Paci E, Puliti D, Carozzi FM, Carrozzi L, Falaschi F, Pegna AL, Mascalchi M, Picozzi G, Pistelli F, Zappa M, Group IW. Prognostic selection and long-term survival analysis to assess overdiagnosis risk in lung cancer screening randomized trials. J Med Screen 2021;28(1):39–47. [PubMed: 32437229]

27. Infante M, Cavuto S, Lutman FR, Passera E, Chiarenza M, Chiesa G, Brambilla G, Angeli E, Aranzulla G, Chiti A, Scorsetti M, Navarria P, et al. Long-term follow-up results of the DANTE trial, a randomized study of lung cancer screening with spiral computed tomography. Am J Respir Crit Care Med 2015;191:1166–75. [PubMed: 25760561]
28. Ebell MH, Lin KW. Accounting for the harms of lung cancer screening. JAMA Intern Med 2018;178:1422–3. [PubMed: 30105383]
29. Pistelli F, Aquilini F, Falaschi F, Puliti D, Ocello C, Lopes Pegna A, Carozzi FM, Picozzi G, Zappa M, Mascalchi M, Paci E, Carrozzi L. Smoking cessation in the ITALUNG lung cancer screening: what does “teachable moment” mean? Nicotine Tob Res 2020;22:1484–91. [PubMed: 31504798]
30. Han SS, Ten Haaf K, Hazelton WD, Munshi VN, Jeon J, Erdogan SA, Johanson C, McMahon PM, Meza R, Kong CY, Feuer EJ, de Koning HJ, et al. The impact of overdiagnosis on the selection of efficient lung cancer screening strategies. Int J Cancer 2017;140:2436–43. [PubMed: 28073150]
31. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger KR, Yatabe Y, Beer DG, Powell CA, Riely GJ, Van Schil PE, Garg K, Austin JH, 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;6:244–85. [PubMed: 21252716]
32. Moyer VA, Force USPST. Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2014;160:330–8. [PubMed: 24378917]
33. Rivera MP, Tanner NT, Silvestri GA, Detterbeck FC, Tammemagi MC, Young RP, Slatore CG, Caverly TJ, Boyd CM, Braithwaite D, Fathi JT, Gould MK, et al. Incorporating coexisting chronic illness into decisions about patient selection for lung cancer screening. An official American Thoracic Society research statement. Am J Respir Crit Care Med 2018;198:e3–e13. [PubMed: 30004250]
34. Caverly TJ, Cao P, Hayward RA, Meza R. Identifying patients for whom lung cancer screening is preference-sensitive: A microsimulation study. Ann Intern Med 2018;169:1–9. [PubMed: 29809244]
35. Iaccarino JM, Steiling KA, Wiener RS. Lung cancer screening in a safety-net hospital: implications of screening a real-world population versus the National Lung Screening Trial. Ann Am Thorac Soc 2018;15:1493–5. [PubMed: 30199643]
36. 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;98:748–56. [PubMed: 16757699]
37. Miller AB, Wall C, Baines CJ, Sun P, To T, Narod SA. Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial. BMJ 2014;348:g366. [PubMed: 24519768]
38. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. Lancet 2012;380:1778–86. [PubMed: 23117178]
39. Ryser MD, Lange J, Inoue LYT, O’Meara ES, Gard C, Miglioretti DL, Bulliard JL, Brouwer AF, Hwang ES, Etzioni RB. Estimation of breast cancer overdiagnosis in a U.S. breast screening cohort. Ann Intern Med 2022;175:471–8. [PubMed: 35226520]
40. Meza R, Jeon J, Toumazis I, Ten Haaf K, Cao P, Bastani M, Han SS, Blom EF, Jonas DE, Feuer EJ, Plevritis SK, de Koning HJ, et al. Evaluation of the benefits and harms of lung cancer screening with low-dose computed tomography: modeling study for the US Preventive Services Task Force. JAMA 2021;325:988–97. [PubMed: 33687469]
41. IARC Working Group on the Evaluation of Cancer-Preventive Interventions. IARC Handbooks of Cancer Prevention Volume 15: Breast cancer screening. Lyon (FR): International Agency for Research on Cancer; 2016.

Novelty and Impact

Across randomised controlled trials of lung cancer screening, 39% of the variation in estimates of overdiagnosis was explained by post-screening follow-up time. Excess incidence declined by 3–4% annually during post-screening follow-up in both the NLST and LUSI trials. The predicted estimate of overdiagnosis with infinite follow-up time was 2% for NLST and 9% for LUSI. Overdiagnosis estimates differed by population characteristics and time trends varied by histological type. This study provides a new perspective on the heterogeneous estimates of overdiagnosis in lung cancer screening. By facilitating a more nuanced understanding of overdiagnosis, the findings could help enable more effective communication with patients about screening benefits and harms.

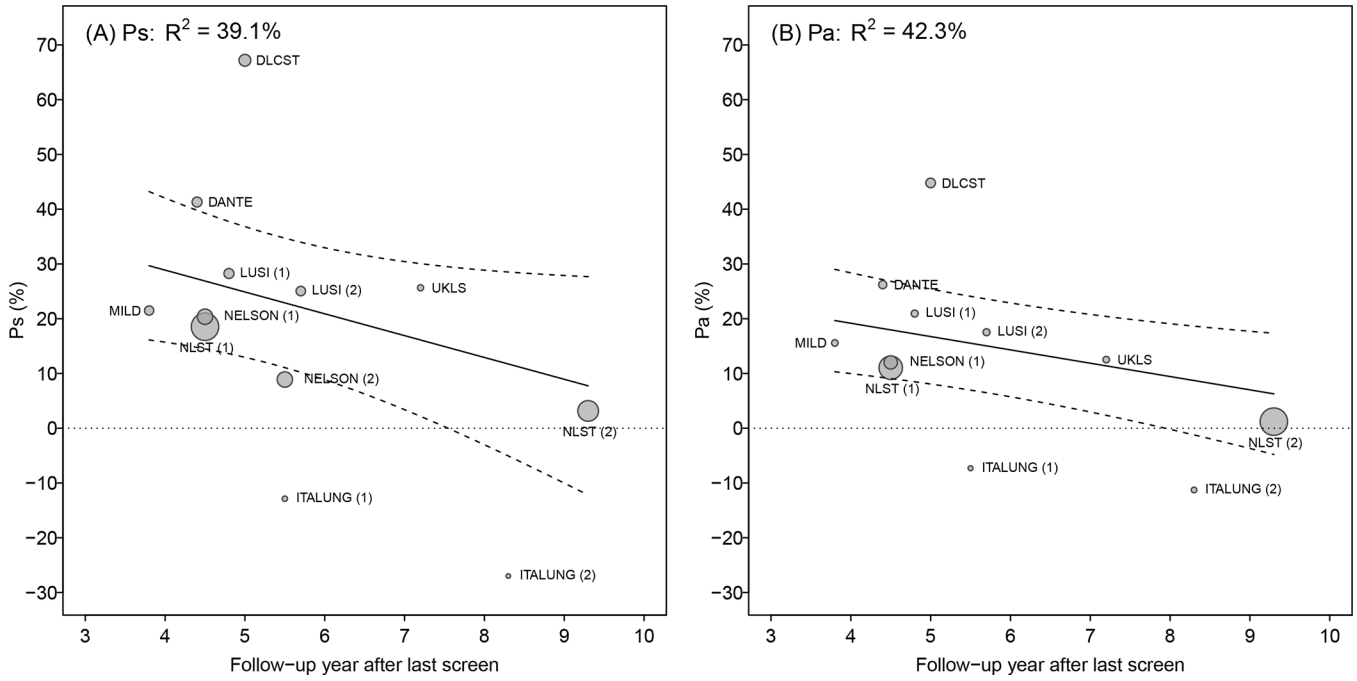

Figure 1: Relationship between (A) excess incidence among screen-detected lung cancers (Ps), (B) excess incidence among lung cancers in the screening arm (Pa), and follow-up time after the last screen in lung cancer screening trials.

The size of each point represents the inverse variance of each study. The solid line denotes the linear regression line fit with multilevel random effects meta regression. The dashed lines represents the confidence intervals.

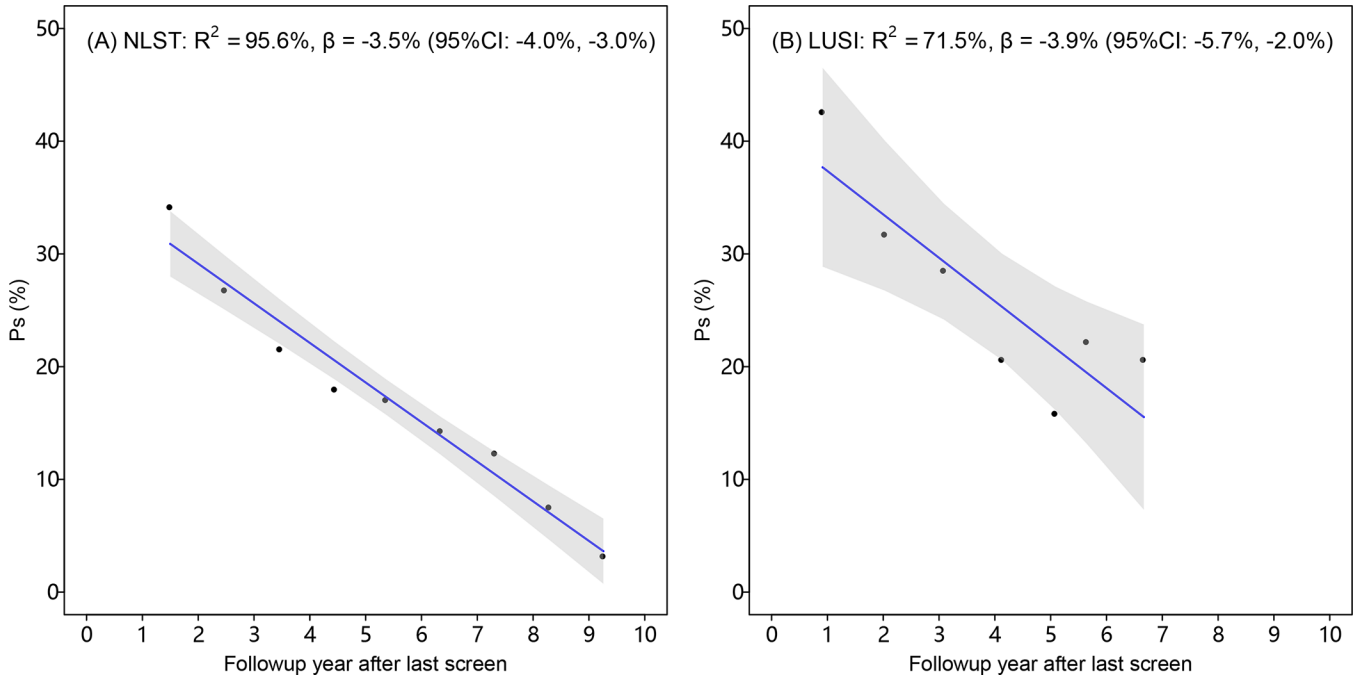

Figure 2: Relationship between excess incidence among screen-detected lung cancers (Ps) and follow-up time after the last screen in the (A) National Lung Screening Trial (NLST) and (B) German Lung Screening Intervention Trial (LUSI).

The blue lines represent the linear regression lines between Ps estimate and follow-up year after last screen; the grey areas show the confidence intervals of the regression lines.

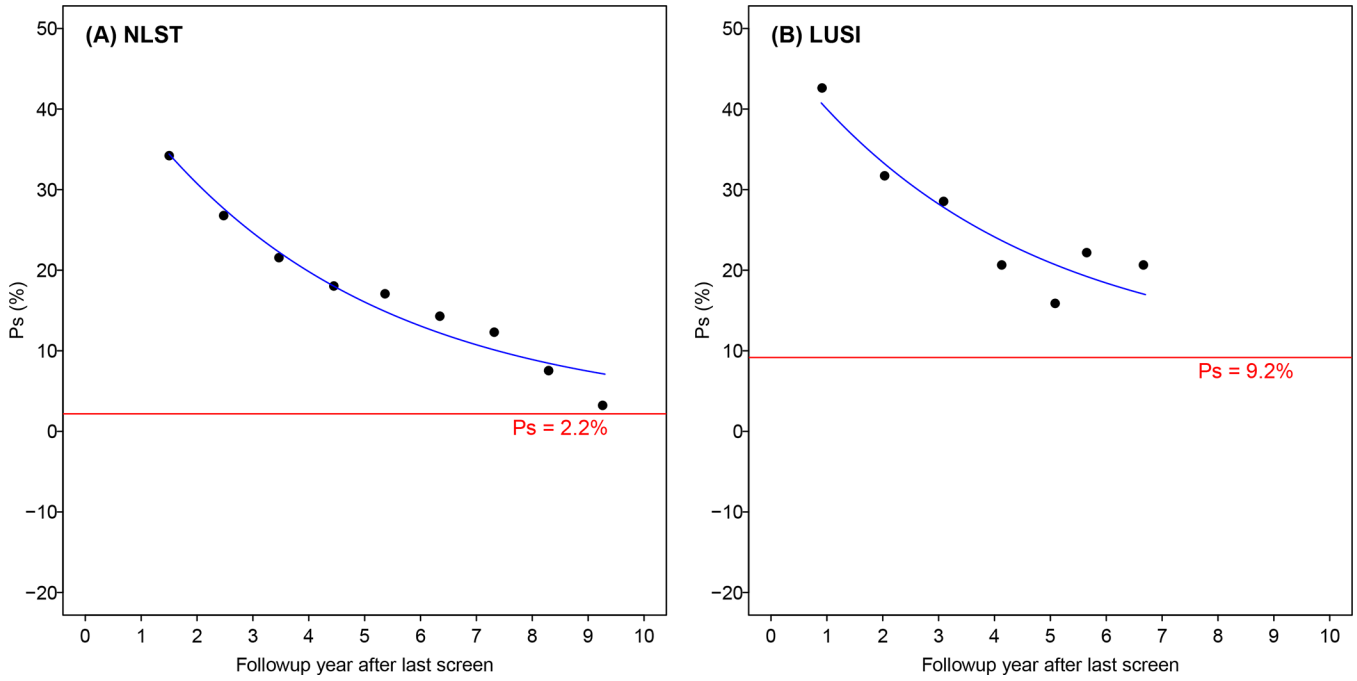

Figure 3: Relationship between excess incidence among screen-detected lung cancer (Ps) and follow-up time after the last screen in (A) National Lung Screening Trial (NLST) and (B) German Lung Screening Intervention Trial (LUSI), based on the exponential decay models with asymptotes.

The blue lines represent the regression lines between Ps estimate and follow-up year after last screen in the exponential decay model; the red lines denote the asymptotes of Ps estimates.

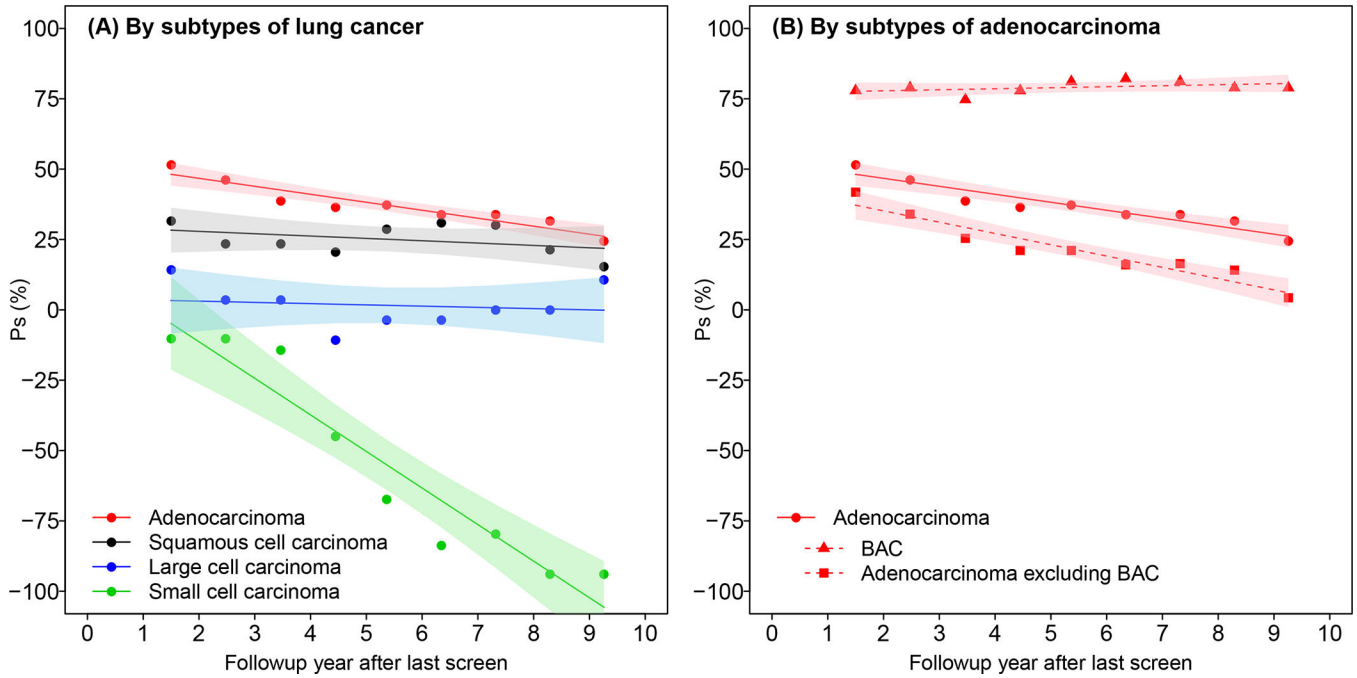

Figure 4: Relationship between excess incidence among screen-detected lung cancers (Ps) and follow-up time after the last screen by histological subtypes of (A) lung cancer overall and (B) adenocarcinoma in the National Lung Screening Trial.

The shaded areas show the confidence intervals of the regression lines. There was statistically significant heterogeneity among the time-trends in Ps across histological types overall (panel A, $P < 0.0001$) and between BAC vs. non-BAC adenocarcinoma (panel B, $P < 0.0001$).

Table 1:

Published and calculated excess lung cancer incidence estimates (Ps and Pa) from randomised controlled trials of low-dose computed tomography (LDCT) lung cancer screening.

Study	Size	Design	Follow-up after last screen, years	Cumulative incidence (screening)	Cumulative incidence (control)	Excess incidence per 1,000	Ps, % [a]	Pa, % [b]
NLST (1) [20, 21]	53,452	3 annual LDCT vs. 3 annual CXR	4.5 (median) [c]	40.8	36.2	4.6	18.5	11.0
NLST (2) [12]	53,452	3 annual LDCT vs. 3 annual CXR	9.3 (median) [c]	63.7	62.9	0.8	3.1	1.2 (calc.)
NELSON (1) [22]	13,195	4 LDCT with 1y, 2y, 2.5y intervals vs. usual care	4.5 (study year)	52.3	46.0	6.3	19.7	12.0 (calc.)
NELSON (2) [22]	13,195	4 LDCT with 1y, 2y, 2.5y intervals vs. usual care	5.5 (study year)	-	-	2.7	8.9	-
DLCST [5, 23]	4,104	5 annual LDCT vs. usual care	5.0 (study year)	46.8	25.8	21.0	67.2	44.8 (calc.)
MILD [24]	4,099	Annual or biennial LDCT vs. usual care (No. rounds not reported)	3.8 (median) [d]	41.2	34.8	6.4	21.5 (calc.)	15.6 (calc.)
LUSI (1) [13]	4,052	5 annual LDCT vs. usual care	4.8 (median) [c]	41.9	33.1	8.8	28.6 (calc.)	20.9 (calc.)
LUSI (2) [9]	4,052	5 annual LDCT vs. usual care	5.7 (median)	44.4	36.6	7.8	25.4	17.8
UKLS[25]	3,968	Single LDCT vs. usual care	7.2 (median)	43.3	37.9	5.4	25.6 (calc.)	12.5 (calc.)
ITALUNG (1) [4]	3,206	4 annual LDCT vs. usual care	5.5 (median) [c]	41.5	44.6	−3.1	−12.9 (calc.)	−7.3 (calc.)
ITALUNG (2) [26]	3,206	4 annual LDCT vs. usual care	8.3 (median) [c]	56.4	62.8	−6.4	−27.0 (calc.)	−11.3 (calc.)
DANTE [27]	2,450	5 annual LDCT vs. yearly clinical review (both groups had baseline CXR and 3-day sputum cytology testing)	4.4 (median) [c]	82.3	60.7	21.6	41.3 (calc.)	26.2 (calc.)

Abbreviations: NLST, National Lung Screening Trial; NELSON, Nederlands–Leuvens Longkanker Screenings Onderzoek; DLCST, Danish Lung Cancer Screening Trial; MILD, Multicentric Italian Lung Detection; LUSI, German Lung Screening Intervention Trial; UKLS, UK lung cancer screening trial; ITALUNG, Italian Lung Cancer Screening Trial; DANTE, Detection And screening of early lung cancer with Novel imaging TEchnology; LDCT, low-dose computed tomography; CXR, chest X-ray.

[a] Ps is calculated as the ratio between the excess incidence and the cumulative incidence of screen-detected lung cancer in the screening arm.

[b] Pa is calculated as the ratio between the excess incidence and the cumulative incidence of all lung cancers in the screening arm.

Note: Estimates were published as-is unless denoted by 'calc.', which indicates that they were calculated based on published numbers of cases or incidence.

[c] Approximate estimates of post-screening follow-up, calculated as the difference between reported median follow-up years and the screening years (the year of the initial round of screening was treated as zero).

[d] Approximate estimates of post-screening follow-up, calculated as the difference between reported follow-up years and median duration of screening.

Table 2:

Excess lung cancer incidence estimates (Ps and Pa) in the National Lung Screening Trial at 9.3 years of median follow-up since the last screen, stratified by baseline characteristics, lung cancer risk, comorbidity, life expectancy, and histological subtype.

Group	Screening group		Control group		Screen-detected cases	Incidence of screen-detected lung cancer (/1000)	Excess incidence of lung cancer (/1000)	Ps		Pa	
	Lung cancer cases	Participants	Lung cancer cases	Participants				% (95%CI) [a]	*P* value	% (95%CI) [b]	*P* value
Overall	1702	26722	1681	26730	649	24.3	0.8	3.3 (−14.2, 19.8)	-	1.3 (−5.5, 7.5)	-
Gender											
Male	999	15769	975	15761	384	24.4	1.5	6.1 (−16.5, 27.1)	0.695	2.4 (−6.4, 10.3)	0.693
Female	703	10953	706	10969	265	24.2	−0.2	−0.7 (−29.2, 24.8)		−0.3 (−10.9, 9.4)	
Age, years											
55–59	468	11440	493	11420	173	15.1	−2.3	−14.9 (−53.0, 18.3)	0.015	−5.5 (−19.4, 6.7)	0.011
60–64	522	8170	519	8199	196	24.0	0.6	2.5 (−30.6, 32.0)		0.9 (−11.5, 11.9)	
65–69	439	4756	421	4761	169	35.5	3.9	10.9 (−23.1, 41.2)		4.2 (−8.8, 15.7)	
70–74	273	2353	248	2344	111	47.2	10.2	21.7 (−18.2, 56.7)		8.8 (−7.4, 22.5)	
Race											
White	1537	24289	1531	24260	599	24.7	0.2	0.7 (−17.5, 17.6)	0.462	0.3 (−6.9, 6.9)	0.545
Non-white	156	2270	144	2263	49	21.6	5.1	23.6 (−50.0, 85.1)		7.4 (−15.2, 25.8)	
Smoking status											
Former	678	13862	622	13830	286	20.6	3.9	19.1 (−5.3, 41.3)	0.108	8.0 (−2.2, 17.4)	0.091
Current	1024	12860	1059	12900	363	28.2	−2.5	−8.7 (−34.0, 14.3)		−3.1 (−12.0, 5.0)	
COPD											
No	1272	21894	1238	21864	486	22.2	1.5	6.6 (−13.4, 25.3)	0.499	2.5 (−5.1, 9.7)	0.499
Yes	423	4674	435	4652	161	34.4	−3.0	−8.7 (−46.3, 23.9)		−3.3 (−17.4, 9.0)	
Lung cancer risk											
≤ 2%	251	9502	244	9518	106	11.2	0.8	7.0 (−37.1, 44.5)	0.966	3.0 (−15.6, 18.6)	0.995

Group	Screening group		Control group		Screen-detected cases	Incidence of screen-detected lung cancer (/1000)	Excess incidence of lung cancer (/1000)	Ps		Pa	
	Lung cancer cases	Participants	Lung cancer cases	Participants				% (95%CI) [a]	*P* value	% (95%CI) [b]	*P* value
2–4%	525	8522	527	8383	184	21.6	−1.3	−5.8 (−42.3, 26.7)		−2.0 (−14.7, 9.3)	
> 4%	791	6728	789	6789	307	45.6	1.4	3.0 (−21.9, 25.6)		1.1 (−8.4, 9.9)	
Charlson comorbidity index [c]											
0	907	16532	885	16476	342	20.7	1.1	5.6 (−19.1, 27.9)		2.1 (−7.2, 10.5)	
1–2	705	9226	713	9241	270	29.3	−0.7	−2.5 (−30.6, 22.6)	0.904	−1.0 (−11.5, 8.6)	0.914
≥ 3	75	687	67	704	32	46.6	14.0	30.1 (−46.3, 90.5)		12.8 (−19.4, 36.5)	
Life expectancy, years [d]											
≤ 10	136	1207	123	1188	57	47.2	9.1	19.4 (−37.6, 66.6)		8.1 (−15.7, 27.1)	
10–20	687	8379	674	8481	252	30.1	2.5	8.4 (−20.1, 34.2)	0.132	3.1 (−7.3, 12.4)	0.098
> 20	666	14152	696	14056	252	17.8	−2.5	−13.8 (−44.5, 13.7)		−5.2 (−16.8, 5.1)	
Histological subtype											
Adenocarcinoma	729	26722	643	26730	351	13.1	3.2	24.6 (4.3, 43.2)		11.8 (2.1, 20.7)	
BAC	121	26722	46	26730	95	3.6	2.8	79.0 (59.6, 96.4)		62.0 (47.6, 73.7)	
Excluding BAC	608	26722	597	26730	256	9.6	0.4	4.4 (−23.1, 29.1)		1.8 (−9.8, 12.2)	
Squamous cell carcinoma	417	26722	396	26730	136	5.1	0.8	15.5 (−27.4, 53.5)	-	5.1 (−8.9, 17.2)	-
Small cell carcinoma	245	26722	291	26730	49	1.8	−1.7	−93.7 (−217.4, −1.6)		−18.7 (−40.8, −0.3)	
Large cell carcinoma	56	26722	53	26730	28	1.0	0.1	10.8 (−79.9, 74.0)		5.4 (−38.0, 35.7)	

Abbreviations: COPD, Chronic obstructive pulmonary disease (including both chronic bronchitis and emphysema); BAC, Bronchiolo-alveolar adenocarcinoma.

[a] Ps is calculated as the ratio between the excess incidence and the cumulative incidence of screen-detected lung cancer.

[b] Pa is calculated as the ratio between the excess incidence and the cumulative incidence of all lung cancers in the screening arm.

[c] Only a few comorbidities (heart disease, stroke, COPD, diabetes, and any cancer) were used to calculate the Charlson comorbidity index due to data availability.

[d] Not all variables needed to build the Life Years Gained From Screening-CT (LYFS-CT) model are available in the NLST trial, i.e., weak/failing kidneys in past year, liver condition in past year, and health problem requiring special equipment. We assumed none of these conditions for all participants. For some variables, there is also no exact information on coronary heart disease, angina pectoris, heart attack, and other heart disease, therefore we used a similar variable "heart disease or heart attack" in the NLST for approximate estimation.