Overdiagnosis in lung cancer screening: Estimates from the German Lung Cancer Screening Intervention Trial

Sandra González Maldonado1,2 | Erna Motsch1,2 | Anke Trotter1 | Hans-Ulrich Kauczor2,3 | Claus-Peter Heusse1,4 | Silke Hermann5 | Sylke Ruth Zeissig6 | Stefan Delorme7 | Rudolf Kaaks1,2

1Division of Cancer Epidemiology (C020), German Cancer Research Center (DKFZ), Heidelberg, Germany
2Translational Lung Research Center Heidelberg (TLRC-H), Member of the German Center for Lung Research, Heidelberg, Germany
3Department of Diagnostic and Interventional Radiology, University Hospital Heidelberg, Heidelberg, Germany
4Department of Diagnostic and Interventional Radiology with Nuclear Medicine, Thoraxklinik Heidelberg GmbH, Heidelberg, Germany
5Epidemiological Cancer Registry Baden-Württemberg, German Cancer Research Center (DKFZ), Heidelberg, Germany
6Cancer Registry of Rhineland-Palatinate, Mainz, Germany
7Department of Radiology, German Cancer Research Center (DKFZ), Heidelberg, Germany

Correspondence
Rudolf Kaaks, German Cancer Research Center, Division of Cancer Epidemiology, Im Neuenheimer Feld 581, 69120 Heidelberg, Germany.
Email: r.kaaks@dkfz-heidelberg.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/ Award Numbers: BE 2486/2-1, BE 2486/2-2

Abstract
Overdiagnosis is a major potential harm of lung cancer screening; knowing its potential magnitude helps to optimize screening eligibility criteria. The German Lung Screening Intervention Trial (“LUSI”) is a randomized trial among 4052 long-term smokers (2622 men), 50.3 to 71.9 years of age from the general population around Heidelberg, Germany, comparing five annual rounds of low-dose computed tomography (n = 2029) with a control arm without intervention (n = 2023). After a median follow-up of 9.77 years postrandomization and 5.73 years since last screening, 74 participants were diagnosed with lung cancer in the control arm and 90 in the screening arm: 69 during the active screening period; of which 63 screen-detected and 6 interval cancers. The excess cumulative incidence in the screening arm (N = 16) represented 25.4% (95% confidence interval: −11.3, 64.3] of screen-detected cancer cases (N = 63). Analyzed by histologic subtype, excess incidence in the screening arm appeared largely driven by adenocarcinomas. Statistical modeling yielded an estimated mean preclinical sojourn time (MPST) of 5.38 (4.76, 5.88) years and a screen-test sensitivity of 81.6 (74.4%, 88.8%) for lung cancer overall, all histologic subtypes combined. Based on modeling, we further estimated that about 48% (47.5% [43.2%, 50.7%]) of screen-detected tumors have a lead time ≥4 years, whereas about 33% (32.8% [28.4%, 36.1%]) have a lead time ≥6 years, 23% (22.6% [18.6%, 25.7%]) ≥8 years, 16% (15.6% [12.2%, 18.3%]) ≥10 years and 11% (10.7% [8.0%, 13.0%]) ≥12 years. The high proportions of tumors with relatively long lead times suggest a major risk of overdiagnosis for individuals with comparatively short remaining life expectancies.

KEYWORDS
lung cancer screening, mean sojourn time, modeling, overdiagnosis, sensitivity

Abbreviations: AIS, adenocarcinoma in situ; BAC, bronchiolo-alveolar carcinoma; DLCST, Danish Lung Cancer Screening Trial; ITALUNG, Italian Lung Cancer Screening Trial; LDCT, low-dose computed tomography; LUSI, Lung Cancer Screening Intervention Trial; MIA, minimally invasive adenocarcinoma; MPST, mean preclinical sojourn time; NELSON, Nederlands Leuvens Longkanker Screening Onderzoek (Dutch-Belgian Randomized Lung Cancer Screening Trial); NLST, US National Lung Screening Trial; NSCLC, non-small cell lung cancer; PA, proportion of all cases; PS, proportion of screen-detected cases.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. International Journal of Cancer published by John Wiley & Sons Ltd on behalf of Union for International Cancer Control.
INTRODUCTION

Randomized trials in the United States and Europe have convincingly shown that screening by low-dose computed tomography (LDCT) is a viable strategy for reducing lung cancer mortality. However, this mortality reduction has to be balanced against various risks of negative side effects, including long-term radiation, risks of surgical interventions of benign lesions after false-positive diagnoses and overdiagnosis.

Overdiagnosis refers to tumors that would not have become manifest in the absence of screening. It is the result of detecting tumors ahead of the time at which they cause symptoms or death—a lead-time window within which a proportion of screening participants may die, before knowing of their lung cancer disease. As the diagnosis of lung cancer generally causes referral to aggressive treatment, overdiagnosis may incur serious and unnecessary losses in quality of life and financial health-care costs. Estimates of its expected magnitude, depending on an individual’s age at the time of screening, may help optimize lung cancer screening eligibility criteria.

In randomized trials, the extent of overdiagnosis has been assessed by determining the excess cumulative incidence in the LDCT screening arm as compared to a control arm without screening, or compared to a control arm using the less sensitive standard chest X-ray as alternative screening method, after the screening had stopped. Initial estimates varied from zero excess in the Italian Lung Cancer Screening Trial (ITALUNG) to 67.2% (95% confidence interval [CI]: 37.1%-95.4%) in the Danish Lung Cancer Screening Trial (DLCST). Conversely, using data from the US National Lung Screening Trial (NLST), the estimated excess incidence rate of LDCT relative to chest X-Rays was 18.5% (5.4%, 30.6%), in the Dutch-Belgian Nederlands Leuvense Longkanker Screening Onderzoek (Dutch-Belgian Randomized Lung Cancer Screening Trial) (NELSON) study, 10 years postrandomization and about 4.5 years since last screening participation the excess incidence estimate among men was 19.7% of screen-detected cases (95% CI: −5.2%-41.6%). In these studies, however, follow-up times after screening cessation varied and likely were too short to cover the longest possible tumor lead times, which is required for unbiased estimation of overdiagnosis in the given population of screening participants.

Further to the excess-incidence method, mathematical modeling has proven useful for generating estimates of overdiagnosis under alternative screening scenarios and follow-up times, beyond those of original trials. Patz et al applied convolution models to data from the NLST to estimate tumor preclinical sojourn time and LDCT screening test sensitivity, and based on these estimates calculated a lifetime excess incidence of 11% of non-small cell lung cancers (NSCLCs) after three annual screenings for a cohort of 60 year old US men and women. Likewise, using a microsimulation model, Schultz et al and ten Haaf et al estimated that, with life-time follow-up, overdiagnosis in the actual NLST trial cohort would shrink to 8.6% of screen-detected cancers (all histologic subtypes combined).

What’s new?

The reduced lung cancer mortality achieved through low-dose computed tomography screening must be balanced against the risk of overdiagnosis. In this randomized screening trial with 5 annual screening rounds and a median post-screening follow-up of 5.73 years, excess lung cancer incidence represented 25.4% of the screen-detected cases. The mean pre-clinical sojourn time was 5.38 years. The probabilities for screen-detected tumors to have remained asymptomatic for 4 and 12 years in the absence of screening were 48% and 11%, respectively. The findings suggest that individuals with short remaining lifetime are at high risk of being overdiagnosed by low-dose computed tomography screening.

MATERIALS AND METHODS

2.1 LUSI trial: Basic design and prospective case ascertainment

The LUSI randomized 50 to 69 year old men and women with a history of heavy smoking (≥25 years of smoking of ≥15 cigarettes per day, or ≥30 years smoking of ≥10 cigarettes per day; ≤10 years since smoking cessation) into a screening intervention arm, comprising a LDCT screening at time of randomization plus four annual follow-up screenings, and a control arm with no intervention. Patients were enrolled between 23 October 2007 and 11 April 2011 and screening was conducted between 23 October 2007 and 25 May 2016.

Besides screen detection, the prospective incidence of lung cancer in both study arms (Supplemental Methods section, Supplemental Table 1) was ascertained comprehensively by a combination of annual follow-up questionnaires (self-reports), and record linkages to cancer registries and mortality registers. For all lung cancer cases, detailed information from medical records (pathology reports, medical letters from responsible physicians on diagnosis and treatment and radiology reports, with their exact dates) was obtained by contacting the treating clinics, and coded to ICD-O-3 for tumor histology and stage (Supplemental Methods section, Supplemental Table 2). Extensive descriptions of study design and results for mortality reduction have been published previously.

2.2 Statistical methods

For the present analyses, performed between February and May 2020, the end-date for follow-up of lung cancer incidence was set at
Excess incidence was calculated as the difference in cumulative incidence of lung cancer in the LDCT and control arms, and expressed as a ratio relative to the cumulative incidence of screen-detected lung cancers (PS). Alternatively, relative to all lung cancers in the LDCT arm (PA) up to 30 April 2019, regardless of whether they were screen-detected or not, 95% CIs were obtained via bootstrapping (5000 repetitions).

Further to calculating excess cumulative incidence, we used a convolution model for clinical incidence assuming a binomial distribution for the observed screen-detected and cases diagnosed either in between screening appointments (interval cancers) or after the last screening appointment, to jointly estimate mean preclinical tumor sojourn time (MPST) and LDCT-based screen detection sensitivity in the screening arm.\textsuperscript{11,12} Based on estimates of MPST, we further

FIGURE 1 Cumulative lung cancer incidence by years from randomization until diagnosis, study arm and tumor histology. After a median follow-up time of 9.77 years postrandomization there were 90 lung cancer cases in the low-dose computed tomography (LDCT) arm, of which 63 were detected by LDCT screening, whereas a total of 74 cases were observed in the control arm. A large proportion of lung cancer cases were adenocarcinomas, with 59 cases (10 of them bronchiolo-alveolar carcinomas [BAC]) observed in the LDCT arm of which 44 were screen-detected, and 37 (1 BAC) in the control arm [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Lung cancer cases by study arm and tumor histology, and empirical estimates of overdiagnosis rates

| Tumor histology          | LDCT | Control | Estimated overdiagnosis rates\(^{a}\) (95% CI) |
|--------------------------|------|---------|---------------------------------------------|
|                          | Screen detected | Nonscreen detected\(^{b}\) | Subtotal | Subtotal | PS | PA |
| All                      | 63   | 27      | 90     | 74       | 25.4 (−11.3, 64.3) | 17.8 (−7.4, 44.7) |
| Adenocarcinoma (%)       | 44   | 15      | 59     | 37       | 50.0 (14.0, 88.4) | 37.3 (11.5, 65.4) |
| BAC (%)                  | 8    | 2       | 10     | 1        | 112.5 (68.2, 113.1) | 90.0 (54.3, 164.4) |
| Non-BAC (%)              | 36   | 13      | 49     | 36       | 36.1 (−8.4, 84.8) | 26.5 (−5.3, 61.8) |
| Other (nonadenocarcinoma) | 19   | 12      | 31     | 37       | −31.6 (−130.8, 83.0) | −19.4 (−76.8, 45.6) |

Abbreviations: BAC, bronchiolo-alveolar carcinoma; LDCT, low-dose computed tomography.
\(^{a}\)Estimated as the excess incidence in the LDCT arm relative to the incidence of screen-detected lung cancers (PS), and alternatively, relative to all lung cancers in the LDCT arm (PA).
\(^{b}\)Nonscreen detected cases are all lung cancer cases diagnosed either between screening rounds or in the follow-up years after the last screening round.
calculated probabilities that screen-detected tumors would have remained in a preclinical (asymptomatic) phase, depending on the follow-up time after screen detection,\textsuperscript{14,15} for lung cancer overall and by subgroups of tumor histology. For details about the statistical methods see the Supplemental Methods section.

Statistical analyses were performed using R, version 3.4.4\textsuperscript{16} and the boot package.\textsuperscript{17}

3 | RESULTS

The 4052 participants in our study (long-term smokers, 2622 males [64.7%]) had a median age of 56.9 (range = 50.3-71.9) years at first screening participation. Counting follow-up time until lung cancer diagnosis, death and date of loss to follow-up or 30 April 2019 as censoring date, whichever came first, there was a median follow-up time of 9.77 (range = 0-11.5, IQR = 8.8-10.4, 10th percentile = 8.2) years since last screening participation. Within this follow-up period, there were 90 lung cancer cases in the LDCT arm, of which 63 were detected by LDCT screening, whereas a total of 74 cases were observed in the control arm (Figure 1). Thus, all histologic subtypes taken together, the excess cumulative incidence until 30 April 2019 amounted to 16 cases, corresponding to more than 16% and 11%, of screen-detected tumors would have remained in a preclinical phase over, respectively, 4, 6, 8, 10 and 12 further years if screening had not been performed.

### TABLE 2

Model parameter estimates and 95% confidence intervals overall and by tumor histology

|             | Mean preclinical sojourn time (years, 95% CI) | Sensitivity (%) (95% CI) |
|-------------|---------------------------------------------|--------------------------|
| Overall     | 5.38 (4.76, 5.88)                           | 81.6 (74.4, 88.8)        |
| Adenocarcinoma | 7.69 (6.49, 8.77)                           | 69.6 (60.8, 79.2)        |
| Adeno BAC   | 8.77 (6.49, 12.20)                          | 100 (92.8, 100)          |
| Adeno non-BAC | 7.69 (6.49, 8.77)                           | 62.4 (53.6, 72.8)        |
| Other (non-adenocarcinoma) | 2.89 (2.49, 3.36)                           | 100 (94.4, 100)          |

Abbreviations: BAC, bronchiolo-alveolar carcinoma; LDCT, low-dose computed tomography.

### TABLE 3

Estimated proportions of screen-detected tumors by lead time \((1 − P_{\text{clin}}(t))\)

| Histologic subtype                | \(1 − P_{\text{clin}}(95\% \text{ CI})\) |
|-----------------------------------|------------------------------------------|
|                                   | 4y                               | 6y       | 8y       | 10y      | 12y      |
| All tumors                        | 47.5% (43.2%, 50.7%)               | 32.8% (28.4%, 36.1%) | 22.6% (18.6%, 25.7%) | 15.6% (12.2%, 18.3%) | 10.7% (8.0%, 13.0%) |
| Adenocarcinoma                    | 59.5% (54.0%, 63.4%)               | 45.8% (39.7%, 50.5%) | 35.3% (29.2%, 40.2%) | 27.3% (21.4%, 32.0%) | 21.0% (15.8%, 25.5%) |
| BAC                               | 63.4% (54.0%, 72.0%)               | 50.5% (39.7%, 61.1%)  | 40.2% (29.2%, 51.9%) | 32.0% (21.4%, 44.0%) | 25.5% (15.8%, 37.4%) |
| Adenocarcinoma, non-BAC           | 59.5% (54.0%, 63.4%)               | 45.8% (39.7%, 50.5%) | 35.3% (29.2%, 40.2%) | 27.3% (21.4%, 32.0%) | 21.0% (15.8%, 25.5%) |
| Other (non-adenocarcinoma)        | 25.1% (20.0%, 30.4%)               | 12.5% (9.0%, 16.7%)    | 6.3% (4.0%, 9.2%)    | 3.1% (1.8%, 5.1%)    | 1.6% (0.8%, 2.8%)    |

Abbreviation: BAC, bronchiolo-alveolar carcinoma.
4 DISCUSSION

In this randomized trial, with a median of 5.73 years (range = 0-11.4, IQR = 4.8-6.3) follow-up after last annual screening by LDCT there was an excess incidence of 24.5% (95% CI: −11.3%, 64.3%) of screen-detected lung cancer cases, all histologic subtypes combined. Expressed as a proportion of screen-detected lung cancer cases, this excess incidence estimate is slightly above that reported for the NELSON study2 (19.7% [95% CI: −5.2%, 41.6%], at about 4.5 years since last screening) or initially reported for the NLST trial8 (18.5% [95% CI: 5.4%-30.6%], at an average follow-up of about 4.5 years after last screening). By contrast, the ITALUNG study3 at 5 years' postscreening follow-up showed no excess (there actually were four fewer cases in the LDCT arm compared to the control arm); PS = −0.11 whereas at approximately 5 years after screening cessation the DLCST7 showed an excess of 67.2% (95% CI: 37.1%-95.4%) of screen-detected cases. Heterogeneity in the excess incidence observed in these independent trials in part may be related to differences in the duration of post-screening follow-up time over which excess incidence was estimated, and to population differences in death rates from competing causes during follow-up (e.g., depending on age, sex, smoking history and other determinants of general health and overall mortality rates). In addition, excess incidence may vary due to differences in the sensitivity of early lung cancer detection, for example, because of heterogeneous protocols used for malignant nodule detection or due to different time intervals between successive screenings (1, 2 and 2.5 years, successively, in the NELSON trial, vs annual screening in the other studies) (Table 4).

Excess incidence is largely caused by lead time due to earlier tumor detection, and does not necessarily reflect overdiagnosis, which exclusively results when lead time exceeds the remaining life expectancy of participants with screen-detected cancers. With increasing duration of follow-up, in randomized stop-screen trials the cumulative incidence gap between screening and control arm generally reduces progressively as even more slowly growing tumors gradually become manifest in the control arm. If follow-up since last screening does not fully cover even the longest detection lead times for all participants in the trial, excess incidence will generally yield an overestimate for overdiagnosis. In our data, the median follow-up time since last screening participation was 5.73 years, and 25% of participants still had follow-up times below 4.8 years. Compared to the mean preclinical sojourn times (MPST), follow-up times in the LUSI trial may have been too short for excess incidence to be a valid estimate for overdiagnosis, and likely this was also the case for reported excess incidence in the DLCST, NELSON and the ITALUNG study so far. The point is well-illustrated by analyses in the NLST trial, which in initial analyses showed an excess incidence of 18.5% after a median follow-up time of 4.5 years after last screening participation,8 whereas in more recent analyses, after an extended period of follow-up to an average of about 9.3 years since last screening, the excess lung cancer incidence in the LDCT had reduced to 3% (compared to the control arm with chest radiography).18 It is important to note, however, that excess incidence in randomized screening trials may provide an overall estimate for the actual populations screened as a whole, but not necessarily for those individuals that would be naturally at highest risk of being overdiagnosed, for example, those with more advanced age, in view of more limited residual life expectancies.

Considering excess incidence as a measure for estimating overdiagnosis, several further factors may cause bias. First, bias may have occurred when screening was also applied in the control arm. Especially in the NLST, where all control arm participants were systematically screened by standard chest radiology (X-ray), excess incidence in the LDCT arm may underrepresent the true magnitude of overdiagnosis by LDCT as compared to no screening at all. In European trials, by contrast, this type of bias may have been minimal as there was no systematic screening in the control arm, and reported rates of screening contamination in the control arms were also low (Table 4). Another potential source of bias is confounding, for example, due to imperfect randomization and resulting imbalances between the trial arms in baseline lung cancer risk, or due to postrandomization differences in factors such as participation in smoking parallel cessation programs that may alter lung cancer risk independently of LDCT screening. For the DLCST, the study investigators reported imperfect randomization, possibly by chance, resulting in significantly more participants with more than 35 pack-years of smoking and a higher proportion of participants with more obstructed lung function in the screening arm as compared to the control arm, which may have contributed to excess lung cancer incidence in the screening arm not caused by LDCT detection.19 In the ITALUNG trial, by contrast, participants in the screening arm were reported to exhibit significantly higher rates of smoking cessation, and lower rates of relapse into smoking among baseline ex-smokers, as compared to usual-care controls.20 which in our study may have actually diminished the excess lung cancer incidence in the screening arm. Finally, independently of bias, excess incidence may vary by tumor histology. Analyses in NLST showed estimates (PS) of 22.5% (95% CI: 9.7%-34.3%) for NSCLC, 78.9% (95% CI: 62.2%-93.5%) for BAC and 11.7% (−3.7%, 25.6%) for NSCLC excluding BAC, and our results show similarly heterogeneous estimates, with high excess incidence especially for all adenocarcinomas (50.0% [14.0%, 88.4%]), and separately for non-BAC adenocarcinomas (36.1% [−8.4%, 84.8%]) and BAC tumors (112.5% [68.2%, 113.1%]).

In our data, although the excess incidence for cancer overall is modest for women, the excess incidence for adenocarcinoma is much larger for women than for men (Supplemental Table 2). We previously reported21 that in LUSI the distribution of histologic subtypes differed significantly between men and women, with women showing a higher proportion of adenocarcinomas, and a much smaller percentage of small cell tumors, than men. Furthermore, LDCT detection (first 5 years after randomization) led to a predominance of diagnosed adenocarcinomas in the screening arm as compared to the control arm, and this was more strongly the case among women than among men. Analyses of data from the NLST and PLCO (Prostate, Lung, Colorectal Ovarian cancer screening) trials in the United States by Ten Haaf et al21 suggested a longer mean preclinical sojourn time (MPST) particularly for adenocarcinoma among women as compared to men, which might explain our observations. Taken together, these observations
| Trial namea | Inclusion criteria | Intervention (screening interval/rounds) | Trial size (intervention/control arm) | Control | FU after last screen (years)b | Participation rate | Contamination rate | Estimated overdiagnosis, PS (95% CI) |
|------------|-------------------|------------------------------------------|---------------------------------------|---------|-----------------------------|-------------------|-------------------|---------------------------------|
| DLCST (Heleno 2018) | Men and women, 50-70 years, ≥20 pack-years (cessation <10 years). | LDCT (12/5) | 4104 (2052/2052) | Usual care | 5 | 95.5% | 20.3% | 0.67 (0.37, 0.96) |
| ITALUNG (Paci 2017, 2020) | Men and women, 55-69 years, ≥20 pack-years in past 10 years | LDCT (12/4) | 3206 (1613/1593) | Usual care | 5 | 81% | Not reported | −0.11 (not reported) |
| LUSI (our study) | Men and women, 50-69 years, ≥25 years of smoking ≥15 cigarettes a day or 30 years of smoking 10 cigarettes per day; cessation <10 years | LDCT (12/5) | 4052 (2029/2023) | Usual care | 5.71d | Lowest at R4 with 93.4% and highest at R1 with 99.9% | 13.0%c (264/2023) | 0.25 (−0.11, 0.64) |
| NELSON (de Koning 2020) | Men, 50-74 years, >15 cigarettes a day for >25 years or >10 cigarettes a day for >30 years; cessation ≤10 years | LDCT (12, 24, 30/4) | 13 195 (6583/6612) | Usual care | 4.5 | 85.8% in total (lowest at R4 with 67.4%, highest at R1 with 95.8%) | Not reported | 0.20 (−0.05, 0.42) |
| NLST (Patz 2014; NLST Team 2019) | Men and women, 55-74 years, ≥30 pack-years, cessation <15 years | LDCT (12/3) or Chest X-Rays (12/3) | 53 452 (26 722/26730) | No control group/CXR | 4.5f 9.3f | 95% | Not reported | 0.19 (0.05, 0.31) |

aResults shown in this table come from studies receiving a quality rating of 1: properly powered and conducted randomized clinical trial.
bReported approximate postscreening follow-up in the original publications.
cApproximate postscreening follow-up calculated as originally reported median post-randomization follow-up = 11.3-3 years corresponding to 4 screening rounds.
dMedian follow-up time after an individual’s last screening appointment and until 30 April 2019.
eContamination rate: percentage of participants in the control group with at least one test (LDCT or X-rays) during the active phase or during follow-up for lung cancer screening purposes outside our screening trial.
fMedian follow-up time after an individual’s last screening appointment as reported by the NLST team.
suggest a greater potential for overdiagnosis in women, a finding further supported by quantitative modeling analyses performed in context of the Cancer Intervention and Surveillance Modeling Network (CISNET). At the same time, our Becker et al. and other data also suggest that, compared to men, LDCT screening may be associated with a greater reduction in lung cancer mortality among women.

Through statistical modeling, we estimated a MPST of 5.38 years (4.76, 5.88), and a screen-test sensitivity of 81.6% (74.4%, 88.8%) for all histologic subtypes combined, and longer MPST especially for non-BAC adenocarcinomas (7.69 years [6.49, 8.77]), combined with 62.4% [53.6%, 72.8%] detection sensitivity) and BAC (8.77 years [6.49, 12.20] with estimated detection sensitivity of 100% [92.8%, 100%]), whereas estimated MPST was shorter for all non-adenocarcinomas (2.89 years [2.49, 3.36], 100% sensitivity [94.4%, 100%]). Our estimates broadly resemble those derived from NLST data by Patz et al., who also reported much longer estimates of MPST for BAC (32.1 [17.3-270.7] years, 38% [7%-62%]) compared to other NSCLC tumors [3.6 [3.00-4.3] years, 83% [72%-94%]). Our results however, are based on substantially smaller numbers of cancer cases, overall and by tumor histology.

The very long MPST for BAC may be particularly noteworthy, BAC, a term previously used for tumors characterized by a lepidic growth pattern, was recently reclassified as a form of lung adenocarcinoma, including adenocarcinoma in situ (AIS), invasive adenocarcinoma and minimally invasive adenocarcinoma (MIA). BAC tumors are characterized by slow growth, are more likely to affect never-smokers, women and young adults, and are associated with significantly better survival than other forms of NSCLC. Based on distinct radiologic appearances, identifying variants of AIS or MIA that require less intensive treatment or just active surveillance could reduce the burden caused by the early detection of these slowly growing tumors.

Based on our MPST estimates we further estimated that, all histologic types combined, 47.5% (43.2%, 50.7%) of screen-detected tumors had a lead time ≥4 years, 32.8% (28.4%, 36.1%) a lead time ≥6 years, and 22.6% (18.6%, 25.7%) a lead time ≥8 years. For individuals whose remaining life expectancies are shorter than these lead times, these different probabilities would correspond to their actual likelihood that a screen-detected tumor would be overdiagnosed. Long-term smokers (a general eligibility criterion for lung cancer screening) may have a relatively low residual life expectancy at age 75 or older. For example, for US men and women age 75 to 79 with a smoking history, remaining lifetime has been estimated at around 7.5 and 10.5 years, respectively. In such population, based on our estimated MPST or assumptions made about the approximate years of life lost due to smoking, they do show how actual overdiagnosis will depend strongly on participants’ age at screen detection and other host factors that may affect residual life expectancy.

Accurate estimation of the parameters underlying the potential magnitude of overdiagnosis is relevant for optimizing lung cancer screening eligibility criteria, in view of maintaining a positive balance between the expected benefit of screening and early tumor detection (ie, life years that may be gained through reduction of lung cancer mortality) vs the likelihood that early tumor detection merely results in overdiagnosis. Taken together, our analyses confirm findings from previous randomized studies that a substantial proportion of lung cancer cases detected by LDCT screening may be overdiagnosed when screening is offered to individuals with comparatively few remaining life years. Some recent analyses suggest that the balance between life years gained vs risk of being overdiagnosed may start becoming less favorable when screening participants reach the age of 75 years. The identification and modeling of further factors, other than age, that may influence the likelihood of overdiagnosis vs expected benefits of LDCT screening, for example, incorporating more refined estimation of residual life expectancy into screening eligibility criteria may further improve guidance for participation in lung cancer screening programs.

ACKNOWLEDGMENTS

We are much obliged to many colleagues who contributed by their engagement to the success of our study. In the years 2007-2011, Marie-Louise Gross, BSc, carried out recruitment with initial patient information and randomization. Kirsten Lenner-Fertig, BSc, worked-up the blood samples and organized the freezer storage. The LUSI trial was designed by Dr. Nikolaus Becker, who coordinated the study as principal investigator from 2007 until his formal retirement in 2018. Andrea Albrecht, BSc, and Ulrike Beckhaus, BSc, mailed the annual questionnaires, performed the scanning of the filled-in questionnaires and data entry into the database, and kept telephone contact in case of doubtful answers or missing feedback. The low-dose computed tomography scans were performed by Jessica Engelhardt, BSc, and Martina Jochim, BSc. Dr. Rudolf Kaaks had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. We would also like to thank Jan Tremper, MD; Monica Eichinger, MD; Daiva Elzbieta Optazaite, MD; Michael Puderbach, MD; and Mark Wielpütz, MD for radiologic evaluations of the low-dose computed tomography images. They were not compensated for their contributions outside their normal salaries. Particular thanks go to all study participants who beautifully complied with the study protocol and thus carried the study to its success. The LUSI study was funded in the years 2007-2010 by the Dietmar Hopp-Foundation together with the German Research Foundation (BE 2486/2-1), and in the years 2010-2013 by the German Research Foundation (BE 2486/2-2). The funding institutions had no involvement in the design of the study, data collection, interpretation of analytic findings or results or the decision to approve publication of the finished manuscript. Open access funding enabled and organized by Projekt DEAL.
CONFLICT OF INTEREST

Dr. Heussel (radiologist) reports research funding, also outside the present study, from Siemens, Pfizer, MeVis Medical Solutions, Boehringer Ingelheim, lectures fees from Gilead Sciences, Essex Pharma, Schering-Plough, AstraZeneca, Eli Lilly and Company, Roche, Merck Sharp & Dohme, Pfizer, Bracco, MEDA Pharma, InterMune, Chiesi Farmaceutici, Siemens, Covidiem, Pierre Fabre, Boehringer Ingelheim, Grifols, Novartis, Basilea, and Bayer and consultation or other fees from Schering-Plough, Pfizer, Basilea, Boehringer Ingelheim, Novartis, Roche, Astellas, Gilead, Merck Sharp & Dohme, Eli Lilly and Company, Intermune, and Fresenius and ownership of stocks from GSK.

The other authors declared no potential conflicts of interest.

ETHICS STATEMENT

LUSI is a clinical research study registered at ISRCTN http://www.isrctn.com/ISRCTN30604390. Ethical approval for the LUSI trial was given by the University of Heidelberg Medical Ethics Committee (073/2001) and by the radiation protection authority (BfS, 22462/2, 2006-045). Written informed consent was provided by all study participants.

DATA AVAILABILITY STATEMENT

Anonymized data related to the participants of the LUSI study and/or parameters derived from their low-dose computed tomography images can be made available, upon reasonable request and under a data transfer agreement approved by the data protection department of the German Cancer Research Center (DKFZ) Heidelberg.

ORCID

Rudolf Kaaks https://orcid.org/0000-0003-3751-3929

REFERENCES

1. Aberle DR, Adams AM, Berg CD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med. 2011; 365:395-409.
2. de Koning HJ, van der Aalst CM, de Jong PA, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. N Engl J Med. 2020;382:503-513.
3. Paci E, Puliti D, Lopes Pegna A, et al. Mortality, survival and incidence rates in the ITALUNG randomised lung cancer screening trial. Thorax. 2017;72:825-831.
4. Pastorino U, Silva M, Sestini S, et al. Prolonged lung cancer screening reduced 10-year mortality in the MILD trial. Ann Oncol. 2019;30:1672.
5. Becker N, Motsch E, Trotter A, et al. Lung cancer mortality reduction by LDCT screening - results from the randomised German LUSI trial. Int J Cancer. 2020;146:1503-1513.
6. Welch HG, Black WC. Overdiagnosis in cancer. J Natl Cancer Inst. 2010;102:605-613.
7. 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;178:1420-1422.
8. Patz EF Jr, Pinsky P, Gatsonis C, et al. Overdiagnosis in low-dose computed tomography screening for lung cancer. JAMA Intern Med. 2014;174:269-274.
9. Schultz FW, Boer R, de Koning HJ. Chapter 7: description of MISCAN-lung. the Erasmus MC lung cancer microsimulation model for evaluating cancer control interventions. Risk Anal. 2012;32 (Suppl 1):S85-S98.
10. Ten Haaf K, de Koning HJ. Overdiagnosis in lung cancer screening: why modelling is essential. J Epidemiol Community Health. 2015;69:1035-1039.
11. Walter SD, Day NE. Estimation of the duration of a pre-clinical disease state using screening data. Am J Epidemiol. 1983;118:865-886.
12. Straatman H, Peer PG, Verbeek AL. Estimating lead time and sensitivity in a screening program without estimating the incidence in the screened group. Biometrics. 1997;53:217-229.
13. Becker N, Motsch E, Gross ML, et al. Randomized study on early detection of lung cancer with MSCT in Germany; study design and results of the first screening round. J Cancer Res Clin Oncol. 2012;138:1475-1486.
14. Paci E, Warwick J, Falini P, Duffy SW. Overdiagnosis in screening: is the increase in breast cancer incidence rates a cause for concern? J Med Screen. 2004;11:23-27.
15. Pashayan N, Duffy SW, Pharoah P, et al. Mean sojourn time, overdiagnosis, and reduction in advanced stage prostate cancer due to screening with PSA: implications of sojourn time on screening. Br J Cancer. 2009;100:1198-1204.
16. R Core Team. The R Package for Statistical Computing: R: A Language and Environment for Statistical Computing. Vienna, Austria: https://www.R-project.org/: R Foundation for Statistical Computing; 2013.
17. Canty A, Ripley B. boot: Bootstrap R (S-Plus) Functions. R package version 1.3-24; 2019.
18. Lung cancer incidence and mortality with extended follow-up in the National Lung Screening Trial. J Thorac Oncol. 2019;14:1732-1742.
19. Wille MM, Dirksen A, Ashraf H, et al. Results of the randomized Danish lung cancer screening trial with focus on high-risk profiling. Am J Respir Crit Care Med. 2016;193:542-551.
20. 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.
21. Ten Haaf K, van Rosmalen J, de Koning HJ. Lung cancer detectability by test, histology, stage, and gender: estimates from the NLST and the PLCO trials. Cancer Epidemiol Biomarkers Prev. 2015;24:154-161.
22. Han SS, Ten Haaf K, Hazeldon WD, et al. The impact of overdiagnosis on the selection of efficient lung cancer screening strategies. Int J Cancer. 2017;140:2436-2443.
23. 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.
24. Gardiner N, Jogai S, Wallis A. The revised lung adenocarcinoma classification-an imaging guide. J Thorac Dis. 2014;6:S537-S546.
25. Boffetta P, Jayaprakash V, Yang P, et al. Tobacco smoking as a risk factor of bronchioloalveolar carcinoma of the lung; pooled analysis of seven case-control studies in the International Lung Cancer Consortium (ILCOCO). Cancer Causes Control. 2011;22:73-79.
26. Ten Haaf K, van der Aalst CM, de Koning HJ. Clinically detected non-aggressive lung cancers: implications for overdiagnosis and overtreatment in lung cancer screening. Thorax. 2018;73:407-408.
27. Østbye T, Taylor DH. The effect of smoking on years of healthy life (YHL) lost among middle-aged and older Americans. Health Serv Res. 2004;39:531-552.
28. Ten Haaf K, Bastani M, Cao P, et al. A comparative modeling analysis of risk-based lung cancer screening strategies. J Natl Cancer Inst. 2020;112:466-479.
29. Katki HA, Cheung LC, Landy R. Basing eligibility for lung cancer screening on individualized risk calculators should save more lives, but life-expectancy matters. J Natl Cancer Inst. 2020;112:429-430.
30. 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;171:623-632.

**SUPPORTING INFORMATION**
Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** González Maldonado S, Motsch E, Trotter A, et al. Overdiagnosis in lung cancer screening: Estimates from the German Lung Cancer Screening Intervention Trial. *Int. J. Cancer*. 2021;148:1097–1105. [https://doi.org/10.1002/ijc.33295](https://doi.org/10.1002/ijc.33295)