Association of computed tomography screening with lung cancer stage shift and survival in the United States: quasi-experimental study

Alexandra L Potter, Allison L Rosenstein, Mathew V Kiang, Shivani A Shah, Henning A Gaisser, David C Chang, Florian J Fintelmann, Chi-Fu Jeffrey Yang

ABSTRACT

OBJECTIVE
To determine the effect of the introduction of low dose computed tomography screening in 2013 on lung cancer stage shift, survival, and disparities in the stage of lung cancer diagnosed in the United States.

DESIGN
Quasi-experimental study using jointpoint modeling, multivariable ordinal logistic regression, and multivariable Cox proportional hazards modeling.

SETTING
US National Cancer Database and Surveillance Epidemiology End Results program database.

PARTICIPANTS
Patients aged 45-80 years diagnosed as having non-small cell lung cancer (NSCLC) between 1 January 2010 and 31 December 2018.

MAIN OUTCOME MEASURES
Annual per cent change in percentage of stage I NSCLC diagnosed among patients aged 45-54 (ineligible for screening) and 55-80 (potentially eligible for screening), median all cause survival, and incidence of NSCLC; multivariable adjusted odds ratios for year-to-year changes in likelihood of having earlier stages of disease at diagnosis and multivariable adjusted hazard ratios for changes in hazard of death before versus after introduction of screening.

RESULTS
The percentage of stage I NSCLC diagnosed among patients aged 55-80 did not significantly increase from 2010 to 2013 (from 27.8% to 29.4%) and then increased at 3.9% (95% confidence interval 3.0% to 4.8%) per year from 2014 to 2018 (from 30.2% to 35.5%). In multivariable adjusted analysis, the increase in the odds per year of a patient having one lung cancer stage lower at diagnosis during the time period from 2014 to 2018 was 6.2% (multivariable adjusted odds ratio 1.062, 95% confidence interval 1.048 to 1.077; P<0.001) higher than the increase in the odds per year from 2010 to 2013. Similarly, the median all cause survival of patients aged 55-80 did not significantly increase from 2010 to 2013 (from 15.8 to 18.1 months), and then increased at 11.9% (8.9% to 15.0%) per year from 2014 to 2018 (from 19.7 to 28.2 months). In multivariable adjusted analysis, the hazard of death decreased significantly faster after 2014 compared with before 2014 (P<0.001). By 2018, stage I NSCLC was the predominant diagnosis among non-Hispanic white people and people living in the highest income or best educated regions. Non-white patients and those living in lower income or less educated regions remained more likely to have stage IV disease at diagnosis. Increases in the detection of early stage disease in the US from 2014 to 2018 led to an estimated 10 100 averted deaths.

CONCLUSIONS
A recent stage shift toward stage I NSCLC coincides with improved survival and the introduction of lung cancer screening. Non-white patients and those living in areas of greater deprivation had lower rates of stage I disease identified, highlighting the need for efforts to increase access to screening in the US.

Introduction
Although advancements in the treatment for late stage lung cancers have been associated with modest improvements in survival, the early detection of lung cancer remains the most promising strategy to decrease mortality from lung cancer. Lung cancer screening using low dose computed tomography has been shown to detect lung cancers at an earlier stage compared with chest radiography or no screening in randomized controlled trials. On the basis of evidence from the National Lung Screening Trial (NLST), the United States Preventive Services Task Force (USPSTF) first recommended low dose computed tomography screening for people at high risk in December 2013. However, skepticism about the effectiveness of such screening outside of rigorous clinical trial settings continues to exist in the medical community. Although screening has been shown to result in a favorable stage shift and improved survival...
in two large scale randomized studies, no national studies have shown these benefits in a real world setting at the population level. The objectives of this study were to evaluate whether the percentage of stage I non-small cell lung cancer (NSCLC) diagnosed increased among patients potentially eligible for screening since 2014, to evaluate whether median all cause survival has increased among these patients since 2014, and to evaluate whether any effects on stage extend to the entire population or only selected parts of the population.

Methods

Data sources
The National Cancer Database (NCDB) is a clinical oncology database jointly administered by the Commission on Cancer of the American College of Surgeons and the American Cancer Society. The NCDB captures approximately 65% of all newly diagnosed cases of lung cancer in the US and Puerto Rico. The NCDB is a hospital based registry and collects data from a different number of participating hospitals each year. Because the number of participating hospitals has been increasing with each year, more recent years include data for a greater number of patients compared with previous years. The Surveillance, Epidemiology, and End Results (SEER) program database is composed of 18 population based US registries that cover approximately 28% of the US population. Data on cancer incidence and mortality are recorded from 2000 to 2018. For the years of study inclusion (2010-18), we reclassified clinical staging by using best available data according to American Joint Committee on Cancer 8th edition criteria.

Study design
This retrospective analysis was approved by the Massachusetts General Hospital Institutional Review Board. We identified patients aged 45-80 in the NCDB and aged 55-79 in the SEER Program database who were diagnosed as having NSCLC from 2010 to 2018 for inclusion by using International Classification of Diseases for Oncology histology and topography codes, 3rd edition (ICD-O-3). We compared patients found to have NSCLC from 1 January 2010 to 31 December 2013 with patients found to have NSCLC from 1 January 2014 to 31 December 2018. We chose these year cut-offs because the USPSTF first recommended low dose computed tomography screening in the US. We also did several analyses to evaluate whether any observed shift in the percentage of stage I disease identified could have been explained by factors other than the start of low dose computed tomography screening in the US. Firstly, we assessed whether changes in the percentage of patients with unknown staging information may have contributed to an increasing percentage of patients having stage I disease at diagnosis (that is, owing to improvements in the accuracy of documenting clinical staging information over time). We also evaluated the percentage of patients with unknown staging information in the NCDB from 2010 to 2018. We also evaluated the percentage of patients with unknown staging information who had surgery and the five year overall survival of patients with unknown staging information. In addition, because most of the patients with unknown staging information were found to have five year overall survival rates similar to patients having stage II and III NSCLC at diagnosis, we did a sensitivity analysis examining the year-to-year percentages of stage I disease identified from 2010 to 2018 among patients identified as having stage I NSCLC from 2010 to 2018 in which we considered patients whom we had previously excluded because they had unknown staging information as having non-indolent histologic subtypes.

Secondly, we compared the per cent change per year in the percentage of NSCLC diagnosed at stage I among patients in the NCDB aged 45-54 (ineligible for screening) with that in patients aged 55-80 (potentially eligible for screening on the basis of age) from 2010 to 2018 to evaluate whether patients who were ineligible for screening also saw an increase in the percentage of stage I lung cancers diagnosed. Thirdly, we assessed whether differential changes occurred in the percentage of patients in the NCDB diagnosed as having stage I NSCLC per year when patients were grouped according to the age of Medicare eligibility: ages 55-64 (ineligible for Medicare) and ages 65-80 (eligible for Medicare).

Fourthly, using the SEER Program database, we compared changes in the year-to-year percentages of stage I disease identified and median all cause survival from 2010 to 2018 among “high screening” states (states with screening rates higher than the national
average lung cancer screening rate in 2018) or “low screening” states (states with screening rates lower than the national average lung cancer screening rate in 2018) (supplementary table B). In addition, using the NCDB, we evaluated differences in the percentages of stage I and stage IV disease diagnosed in 2010 versus 2018 according to patients’ race/ethnicity and area of residence (high income versus low income and well educated versus less educated) (see supplementary methods).

Lastly, we evaluated trends in incidence of lung cancer by stage from 2010 to 2018 using the SEER Program database. We also estimated the number of deaths from NSCLC averted owing to a shift toward earlier stages of disease identified after the introduction of lung cancer screening in the US (see supplementary methods).

Statistical analysis
We evaluated trends from 2010 to 2018 in the percentage of patients with stage I NSCLC at diagnosis and in the median all cause survival among patients in the study cohort by using joinpoint regression (see supplementary methods). We also used joinpoint regression to characterize trends in incidence of lung cancer (see supplementary methods) by stage group in the SEER Program database. We did sensitivity analyses allowing for up to two joinpoints.

In addition, using the NCDB, we investigated changes in the likelihood of having a lower stage (for example, stage I instead of stage II, stage II instead of stage III, or stage III instead of stage IV) at diagnosis of NSCLC from 2010 to 2018 at an individual level by using a multivariable ordinal logistic regression model. Clinical stage at diagnosis (stage I, II, III, or IV) was the outcome, and we included the introduction of low dose computed tomography lung cancer screening in the US as a time varying covariate. We adjusted for sex, age, race/ethnicity, Charlson-Deyo comorbidity score, median census tract income, percentage of people without a high school education living in that patients’ zip code, insurance status, distance from hospital, histologic subtype, tumor grade, facility type, metropolitan/urban/rural status, and region of residence (for example, North East, East North Central, Pacific). Standard errors were clustered at the hospital level. We did sensitivity analyses varying the time of the introduction of lung cancer screening in the US.

We also did stratified interaction tests with the introduction of low dose computed tomography screening included as a time varying covariate. We first stratified patients according to the age of eligibility for lung cancer screening: patients aged 45-54 (ineligible for screening) and patients aged 55-80 (potentially eligible for screening on the basis of age). Additionally, we stratified patients who were potentially eligible for screening according to the age of Medicare eligibility: patients aged 55-64 (potentially eligible for screening and ineligible for Medicare) and patients aged 65-80 (potentially eligible for screening and eligible for Medicare).

In addition, we used a multivariable Cox proportional hazards model to evaluate changes in the hazard of death before and after the USPSTF recommended low dose computed tomography screening in December 2013. We adjusted for the same covariates listed above in addition to clinical stage group and the receipt of surgery, chemotherapy, radiation, and immunotherapy. We estimated the number of deaths averted due to a shift toward earlier stages of disease identified on the basis of the hazard function from the multivariable Cox proportional hazards model (see supplementary methods). We used Joinpoint Software Version 4.9.0.0 and Stata/MMP Version 17.0 for PC for statistical analyses.

Patient and public involvement
Although we did not directly involve patients or the public in the design, reporting, or dissemination plans of this research, our experiences speaking to patients and members of the public through lung cancer screening awareness and advocacy efforts with the American Lung Cancer Screening Initiative inspired this study. We did not ask patients or members of the public to advise on the interpretation or writing up of results.

Results
Between 2010 and 2018, 763474 patients in the NCDB met the study inclusion criteria (supplementary figure A). Table 1 shows the characteristics of patients in the NCDB diagnosed as having NSCLC from 2010 to 2013 and from 2014 to 2018.

Lung cancer stage shift since USPSTF first recommended screening in December 2013
Between 2010 and 2018, the percentage of patients aged 55-80 in the NCDB with stage I NSCLC (as opposed to other stages of NSCLC) at diagnosis increased by 28% (from 27.8% to 35.5%; P<0.001). From 2010 to 2013, the year-to-year percentages of patients with stage I NSCLC at diagnosis did not increase significantly (from 27.8% to 29.4%; annual per cent change 1.9%, 95% confidence interval −0.1% to 3.8%; P=0.06) (fig 1). However, an inflection point was identified corresponding to an accelerated increase in the rate of stage I disease diagnosed in 2013 (95% confidence interval 2012 to 2015). From 2014 to 2018, the year-to-year percentages of stage I disease diagnosed increased at more than twice the rate seen in the preceding four year term (from 30.2% to 35.5%; annual per cent change 3.9%, 3.0% to 4.8%; P<0.001).

In multivariable adjusted analysis, from 2010 to 2013 the increase in the odds of a patient being diagnosed as having one lung cancer stage lower (for example, stage I instead of stage II, stage II instead of stage III, or stage III instead of stage IV) was 2.3% higher (multivariable adjusted odds ratio 1.023, 95% confidence interval 1.013 to 1.033; P<0.001) compared with the previous year. By contrast, from 2014 to 2018 the increase in the odds of a patient being diagnosed as having one lung cancer stage lower was 8.7%
We compared the year-to-year changes in the percentage of stage I NSCLC diagnosed from 2010 to 2018 among patients aged 45-54 (ineligible for screening) with those among patients aged 55-64 (potentially eligible for screening and ineligible for Medicare) and aged 65-80 (potentially eligible for screening and eligible for Medicare). From 2010 to 2013, the rate of stage I disease identified among patients aged 45-54 increased gradually while it was relatively constant among patients aged 55-64 and higher compared with the previous year. The increase in the odds per year of a patient being diagnosed as having one lung cancer stage lower from 2014 to 2018 was 6.2% (multivariable adjusted odds ratio 1.062, 1.048 to 1.077; P<0.001) higher than the increase in the odds per year of a patient being diagnosed as having one stage lower from 2010 to 2013. Sensitivity analyses varying the year of the introduction of lung cancer screening confirmed that the inflection point corresponding to an accelerated shift toward earlier stages of disease identified occurred in 2014 (supplementary figure B).

Changes in diagnosis of indolent lung cancers from 2010 to 2018

We did an analysis of patients diagnosed as having stage I NSCLC between 2010 and 2018 by histologic subtype. From 2010 to 2018, the percentage of stage I NSCLC diagnosed decreased or did not change (supplementary table C). Furthermore, in an analysis of only patients found to have non-indolent histologic subtypes, an inflection point was identified corresponding to an accelerated increase in the rate of stage I disease diagnosed in 2013 (2012 to 2015) (supplementary figure C).

Changes in percentage of patients with unknown staging information from 2010 to 2018

The percentage of patients in our cohort with unknown staging information decreased from 12.2% in 2010 to 9.1% in 2018 (supplementary figure D). Compared with patients with stage I disease, patients with unknown staging information were less likely to have surgery (40% v 72%; P<0.001) and had significantly worse five year all cause survival (34.1% v 59.5%, log-rank P<0.001). We did a sensitivity analysis of the year-to-year percentages of patients with unknown staging information from 2010 to 2018 in which we included patients with known staging information and patients with unknown staging information. Patients with unknown staging information were categorized as having disease more advanced than stage I NSCLC (on basis of treatment and survival data reported above). In this sensitivity analysis, an inflection point corresponding to an accelerated increase in the rate of stage I disease diagnosed was identified in 2012 (2012 to 2015) (supplementary figure E).

Lung cancer stage shift among people potentially eligible versus ineligible for screening

We compared the year-to-year changes in the percentage of stage I NSCLC diagnosed from 2010 to 2018 among patients aged 45-54 (ineligible for screening) with those among patients aged 55-64 (potentially eligible for screening and ineligible for Medicare) and aged 65-80 (potentially eligible for screening and eligible for Medicare). From 2010 to 2013, the rate of stage I disease identified among patients aged 45-54 increased gradually while it was relatively constant among patients aged 55-64 and higher compared with the previous year. The increase in the odds per year of a patient being diagnosed as having one lung cancer stage lower from 2014 to 2018 was 6.2% (multivariable adjusted odds ratio 1.062, 1.048 to 1.077; P<0.001) higher than the increase in the odds per year of a patient being diagnosed as having one stage lower from 2010 to 2013. Sensitivity analyses varying the year of the introduction of lung cancer screening confirmed that the inflection point corresponding to an accelerated shift toward earlier stages of disease identified occurred in 2014 (supplementary figure B).
increased slightly among patients aged 65-80 (fig 1). By contrast, inflection points were identified in 2013 (2012 to 2013) and 2014 (2012 to 2016) corresponding to accelerated increases in the percentage of stage I disease diagnosed among patients aged 55-64 and 65-80, respectively; no joinpoint corresponding to an accelerated increase in the rate of stage I disease diagnosed was identified among patients aged 45-54 from 2010 to 2018.

We also did a stratified interaction test with the introduction of lung cancer screening included as a time varying covariate among patients who were potentially eligible for screening (age 55-80) and patients who were ineligible for screening (age 45-54). Among patients who were ineligible for lung cancer screening, the increase in the odds of having one lung cancer stage lower at diagnosis per year from 2014 to 2018 was not significantly greater compared with the increase in the odds of having one lung cancer stage lower at diagnosis per year from 2010 to 2013. The test for the interaction was not significant (adjusted odds ratio 1.000002, 0.9999903 to 1.000014; P=0.71).

**Rates of increase in survival of patients with lung cancer since December 2013**

Between 2010 and 2013, the median all cause survival of patients aged 55-80 diagnosed as having NSCLC increased from 16 months in 2010 to 18.3 months in 2013 (annual per cent change 4.1%, −0.3% to 8.9%; P=0.06) (fig 2). A jointpoint was identified in 2013 (2012 to 2015). From 2014 to 2018, the rate of increase in the median all cause survival of patients aged 55-80 diagnosed as having NSCLC nearly tripled (annual per cent change 11.9%, 8.9% to 15.0%; P=0.001). By 2018, the median all cause survival was 28.2 (95% confidence interval 27.7 to 28.6) months.

In multivariable adjusted analysis, the hazard of death decreased significantly faster after 2014 than before 2014 (multivariable adjusted hazards ratio 0.963, 0.954 to 0.972; P<0.001). We estimate that improvements in survival due to a shift toward earlier stages of disease identified from 2014 to 2018 translates into an estimated 10100 deaths averted (supplementary figure F).

**Percentage of stage I disease identified and survival of patients with lung cancer in high screening states versus low screening states**

We compared the rate of stage I disease identified per year from 2010 to 2018 among patients living in high screening states and low screening states. The year-to-year percentages of stage I disease identified did not significantly change among patients living in high screening states from 2010 to 2013 (annual per cent change 0.9%, −1.9% to 3.8%; P=0.41) or among patients living in low screening states from 2010 to 2012 (1.0%, −2.0% to 4.1%; P=0.40) (fig 3, top panel). However, from 2014 to 2018, the percentage of stage I disease identified in high screening states increased significantly faster compared with low screening states (annual per cent change 5.4% (4.1% to 6.7%) versus 3.7% (3.2% to 4.2%)).

From 2010 to 2014, the median all cause survival of patients living in high screening states did not change (annual per cent change 2.2%, −4.7% to 9.7%; P=0.39) (fig 3, bottom panel). A a joinpoint was identified in 2014 (2012 to 2015). From 2015 to 2017, the median all cause survival of patients living in high screening states accelerated (annual per cent change 13.4%, 1.4% to 26.9%; P=0.04). By contrast, no inflection point was identified corresponding to an accelerated increase in the median all cause survival of patients living in low screening states.

**Table 1 | Continued**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>2010-13 (n=314107)</th>
<th>2014-18 (n=449367)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Histology</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>153266 (48.8)</td>
<td>265321 (54.6)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>88400 (28.1)</td>
<td>126441 (28.1)</td>
<td></td>
</tr>
<tr>
<td>Large cell carcinoma</td>
<td>7076 (2.3)</td>
<td>7400 (1.6)</td>
<td></td>
</tr>
<tr>
<td>Adenosquamous cell carcinoma</td>
<td>5071 (1.6)</td>
<td>6402 (1.4)</td>
<td></td>
</tr>
<tr>
<td>Carcinoid tumor</td>
<td>11211 (3.6)</td>
<td>17563 (3.9)</td>
<td></td>
</tr>
<tr>
<td>Tumors formerly classified as bronchioloalveolar carcinoma</td>
<td>13558 (4.3)</td>
<td>14755 (3.3)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>53523 (11.3)</td>
<td>31485 (7.0)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Grade</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well differentiated</td>
<td>22152 (7.1)</td>
<td>32382 (7.2)</td>
<td></td>
</tr>
<tr>
<td>Moderately differentiated</td>
<td>68082 (21.7)</td>
<td>96642 (21.5)</td>
<td></td>
</tr>
<tr>
<td>Poorly differentiated</td>
<td>91241 (29.7)</td>
<td>110963 (24.7)</td>
<td></td>
</tr>
<tr>
<td>Undifferentiated</td>
<td>3946 (1.3)</td>
<td>3881 (0.9)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>126686 (40.3)</td>
<td>205499 (45.7)</td>
<td></td>
</tr>
<tr>
<td><strong>Year of diagnosis:</strong></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>2010</td>
<td>75329 (24.0)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>76713 (24.4)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>79157 (25.2)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>82908 (26.4)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>0 (0)</td>
<td>86650 (19.3)</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>0 (0)</td>
<td>89239 (19.9)</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>0 (0)</td>
<td>92081 (20.5)</td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>0 (0)</td>
<td>94928 (21.3)</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>0 (0)</td>
<td>86469 (19.2)</td>
<td></td>
</tr>
</tbody>
</table>

*Average household income of zip code where patient lives. $1=£0.75 or €0.92.
†Proportion of adults aged ≥25 in patient’s zip code who did not graduate from high school.
‡Distance between patient’s residence (based on zip code or city centroid) and hospital at which patient received diagnosis and/or was treated.
Disparities in stage of lung cancer identified among certain patient populations

Figure 4 shows the percentage of cases of NSCLC diagnosed at stage I and at stage IV by patient’s race and ethnicity, by median income of the patient’s area of residence, and by percentage of people with no high school education living in the patient’s area of residence in 2010 versus 2018. In 2010, stage IV disease was the predominant stage diagnosed among patients of all races and ethnicities who were aged 55-80. However, in 2018, for the first time, more non-Hispanic white patients had stage I disease than stage IV disease at diagnosis (36.9% v 36.2%). In comparison, stage IV disease continued to be the predominant stage diagnosed among non-white patients in 2018. Furthermore, in 2018, stage I had become the predominant diagnosis among patients with NSCLC aged 55-80 living in the highest income and best educated regions. Patients who lived in lower income or less educated regions remained more likely to have stage IV disease at diagnosis in 2018.

Incidence of early stage NSCLC diagnosed among people aged 55-79

From 2010 to 2013, the number of cases of stage I NSCLC diagnosed in the US among people aged 55-79 increased slightly (annual per cent change 3.4%, 0.5% to 6.3%; P=0.03) (fig 5). A joinpoint was identified in 2013 (2012 to 2015). From 2014 to 2018, the number of cases of stage IV NSCLC diagnosed increased slightly between 2010 and 2016 (annual per cent change 1.8%, P=0.30; 55-64 (2010-13): 1.3%, P=0.06; 55-80 (2010-13): 1.3%, P=0.06; 55-80 (2013-18): 3.9%, P=0.001; 45-54 (2010-18): 3.4% P=0.001; 55-64 (2010-13): 1.3%, P=0.001; 55-64 (2013-18): 4.9%, P=0.001; 65-80 (2010-14): 2.3%, P=0.001; 65-80 (2014-18): 3.8%, P=0.001. “Raw” denotes actual percentage of patients with stage I disease at diagnosis each year. “Model” denotes results of joinpoint regression modeling. An interactive version of this graphic is available at https://public.flourish.studio/visualisation/8864567/

Discussion

In this national analysis, we found that the percentage of stage I NSCLC diagnosed among patients potentially
eligible for screening has increased significantly year over year starting in 2014, shortly after the USPSTF first recommended screening in December 2013. Median all cause survival of patients with NSCLC who were potentially eligible for screening also increased significantly starting in 2014. The annual rate of increase in the percentage of stage I disease identified and the median all cause survival of patients diagnosed as having NSCLC increased significantly faster from 2014 to 2018 among patients living in high screening states compared with those living in low screening states. We found disparities in the stage of lung cancer at diagnosis among subgroups of patients. Patients who were non-Hispanic white and patients living in better educated and higher income areas were more likely to have stage I disease than stage IV disease at diagnosis. Increases in the diagnosis of early stage disease in the US led to an estimated 10,100 averted deaths due to earlier detection of NSCLC from 2014 to 2018.

**Possible explanations for findings**

Although we found a clear shift to stage I NSCLC identified among patients aged 55-80 in our cohort from 2014 to 2018, corresponding to the advent of low dose computed tomography screening in the US, we assessed whether other reasons could have explained the observed stage shift. Firstly, we evaluated whether the observed lung cancer stage shift in the US from 2010 to 2018 was accompanied by a more frequent detection of lung tumors of “indolent” histologic subtypes (for example, typical carcinoid and subtypes formerly classified as bronchioloalveolar carcinoma). One concern associated with lung cancer screening is the risk of detecting indolent disease that may never clinically manifest. However, we found that the accelerated increase in the percentage of stage I NSCLC identified (as opposed to other stages), which started in 2014, was not accompanied by an increase in the percentage of indolent lung cancers diagnosed. Furthermore, in a sensitivity analysis excluding indolent histologic subtypes, an inflection point was still identified corresponding to an accelerated increase in the rate of stage I disease diagnosed. These findings suggest that overdiagnosis of lung cancer likely did not contribute significantly to the accelerated increase in the rate of stage I disease identified beginning in 2014.

Secondly, we examined whether decreases in the percentage of patients with unknown staging data could have contributed to increases in the rate of stage I disease diagnosed from 2010 to 2018. We found that the percentage of patients with unknown staging data decreased from 2010 to 2018. However, patients in our cohort with unknown staging data were significantly less likely to have surgery and had significantly worse five year all cause survival than patients with stage I
Fig 3 | Trends from 2010 to 2018 in percentage of stage I non-small cell lung cancer diagnosed (top) and in median all cause survival (bottom) among patients aged 55-79 in Surveillance Epidemiology End Results Program database living in high screening (US states where lung cancer screening rate is above national average lung cancer screening rate in 2018) and low screening (US states where lung cancer screening rate is below national average lung cancer screening rate in 2018) states. Each consecutive annual interval is represented as a point value. Vertical dotted line represents date when United States Preventive Services Task Force (USPSTF) first recommended low dose computed tomography (LDCT) lung cancer screening. Annual percentage changes were as follows. Low screen, percentage diagnosed 2010-12: 1.0%, P=0.40; 2012-18: 3.7%, P<0.001. High screen, percentage diagnosed 2010-13: 0.9%, P=0.41; 2013-18: 5.4%, P<0.001. Low screen, median survival 2010-17: 5.9%, P<0.001. High screen, median survival 2010-14: 2.2%, P=0.39; 2014-17: 13.4%, P=0.04. "Raw" denotes actual percentage of patients with stage I disease at diagnosis and median all cause survival of patients each year. "Model" denotes results of joinpoint regression modeling. An interactive version of this graphic is available at https://public.flourish.studio/visualisation/8876566/
Fig 4 | Percentage of cases of non-small cell lung cancer (NSCLC) diagnosed in National Cancer Database at stage I and stage IV in 2010 versus 2018 by patient’s race/ethnicity, median income of people living in patient’s area of residence, and percentage of people with no high school education living in patient’s area of residence (see supplementary methods for details). Purple bars represent percentage of NSCLC cases diagnosed at stage I. Yellow bars represent percentage of NSCLC cases diagnosed at stage IV. Median income is grouped into low income (<$40,227), middle-low income ($40,227-$50,353), middle-high income ($50,354-$63,332), and high income (>63,332). Percentage of people with no high school education living in patient’s area of residence is grouped into low education (≥17.6%), middle-low education (10.9-17.5%), middle-high education (6.3-10.8%), and high education (<6.3%). An interactive version of this graphic is available at https://public.flourish.studio/visualisation/8877224/
disease at diagnosis, suggesting that these patients likely had more advanced stages of disease.

Thirdly, we evaluated whether an inflection point existed corresponding to an acceleration in the percentage of stage I disease identified among patients aged 45-54 (who were ineligible for screening) from 2014 to 2018. Over the past two decades, the use of computed tomography coronary angiography and chest imaging has been increasing in hospital settings.14 15 Previous studies have shown that this increase in diagnostic imaging has been accompanied by a growing number of incidental pulmonary nodules identified.15 16 The increased detection of incidental pulmonary nodules may have affected the rate of diagnosis of early stage lung cancer. If increased detection of incidental pulmonary nodules were responsible for an accelerated increase in the percentage of stage I disease identified among patients aged 55-80 in our cohort, we would expect to find a similar inflection point in 2014 among patients aged 45-54 whereas we identified inflection points among patients aged 55-64 and 65-80 (age groups that match the USPSTF screening criteria). Although we recognize that patients aged 45-54 are less likely to have comorbid disease and therefore may undergo diagnostic imaging less frequently than older patients, we would still expect to observe an inflection point among patients aged 45-54 if an increasing use of non-screening related diagnostic imaging was driving the accelerated increase in the percentage of stage I disease identified. Therefore, an increase in the use of non-screening related diagnostic imaging alone is unlikely to have resulted in the sharp increase in the percentage of stage I disease identified among patients aged 55-80 in the US from 2014 to 2018.

The rate of diagnosis of stage I disease increased from 2010 to 2013, before low dose computed tomography screening was first recommended by the USPSTF. Although we cannot be certain why this increase occurred, it may be partially explained by some efforts that were underway before 2013 to increase early detection of lung cancer through the
more regular use of low dose computed tomography or chest radiography. An increasing use of non-screening diagnostic imaging in hospital settings and changes in the population at risk over time may also have contributed to increases in the rate of stage I disease identified from 2010 to 2013. However, it is unlikely that any one of these possible explanations would result in an inflection point identified in 2013 corresponding to an accelerated increase in the rate of stage I disease diagnosed from 2014 to 2018.

Fourthly, we considered the possibility that concomitant policy changes during the study period could affect the year-to-year changes in the rate of stage I disease identified. For example, major provisions of the Affordable Care Act were implemented in the US in January 2014, resulting in millions of previously uninsured Americans gaining access to health insurance. Large increases in the number of Americans with health insurance could contribute to a shift toward earlier stages of disease identified as a result of improved access to medical care. However, implementation of the Affordable Care Act would likely have had minimal effects on Americans over age 65, given that the large majority of these people were already insured by Medicare. If the adoption of provisions of the Affordable Care Act drove the accelerated increase in the rate of stage I disease identified beginning in 2014, we would not have expected to see an inflection point corresponding to an accelerated rate of increase of stage I disease diagnosed among patients aged 65-80 (who were likely unaffected by the provisions implemented by the Affordable Care Act in 2014). However, we still identified an inflection point among this age group, suggesting that the implementation of major provisions of the Affordable Care Act does not fully explain the accelerated rate of increase of stage I disease identified beginning in 2014.

Fifthly, we evaluated whether the year-to-year changes in the percentage of stage I disease identified differed between high screening states and low screening states. If the accelerated increase in the rate of stage I disease identified was explained by factors other than the start of screening in the US, we would expect to identify similar increases in the rate of stage I disease diagnosed among people living in US states reporting lung cancer screening rates above and below the national average screening rate. However, we found that the rate of stage I disease identified increased significantly faster from 2014 to 2018 among high screening states than among low screening states. Taken together, these findings suggest that the accelerated increase in the rate of stage I disease identified is likely explained by the start of lung cancer screening in the US.

In addition, we found that the annual per cent increase in the median survival of patients with lung cancer aged 55-80 in the study cohort tripled from 2014 to 2018 compared with 2010 to 2013. This is likely due to the increasing incidence of stage I disease diagnosed and decreasing incidence of stage III disease and relatively stable incidence of stage IV disease diagnosed (a slight increase (from 2014 to 2016) and decrease (from 2017 to 2018) occurred in the incidence of stage IV disease). Of note, two previous studies using the SEER Program database have reported that recent increases in the incidence of stage I NSCLC diagnosed contributed to improvements in lung cancer survival.

These studies similarly suggested that these trends may be explained, in part, by increases in the early detection of lung cancer through screening. Our analysis differed from these studies in that we focused on evaluating patients aged 55-80 (patients potentially eligible for screening on the basis of age) and we used the NCDB to evaluate whether the rate of increase in stage I disease was significantly greater after the USPSTF recommended screening than it was before.

Our findings also highlight disparities between certain patient populations in the stage of NSCLC identified. In 2018, for the first time ever, stage I disease was the predominant stage of NSCLC identified among non-Hispanic white patients. By contrast, stage IV disease remained the predominant stage of NSCLC diagnosed among non-white patients in 2018. Furthermore, by 2018 stage I was the predominant stage of NSCLC diagnosed among patients who lived in the highest income and best educated regions, whereas patients living in lower income and less educated regions continued to be predominantly diagnosed as having stage IV NSCLC. These differences likely reflect disparities in access to screening between certain populations in addition to the poor adoption of low dose computed tomography screening in community settings. Although lung cancer screening is recommended for people meeting requisite age and smoking history criteria, many challenges that prevent the widespread adoption of screening in the US remain. These challenges, which can include the cost of screening for people who are underinsured or uninsured, limited access to lung cancer screening centers, and low awareness of lung cancer screening, disproportionately affect certain communities, including rural communities, lower income communities, and less educated communities. Furthermore, previous studies have shown that the USPSTF lung cancer screening guidelines preclude screening of many women and racial minorities at high risk, given their lower pack year smoking history and earlier age at diagnosis of lung cancer compared with white men diagnosed as having lung cancer. Differences in the percentage of people at high risk meeting screening eligibility criteria among certain populations and the inability to support lung cancer screening programs in community settings may contribute to widening disparities in the early diagnosis and survival of patients with lung cancer in the near future if efforts to increase eligibility for screening and to increase adoption of screening are neglected.

The disparities reported in this study likely also reflect broader structural inequalities whereby many patients lack access to high quality healthcare.
The COVID-19 pandemic has likely further exacerbated these disparities in the screening for and early detection of lung cancer. As we recover from the pandemic, coordinated awareness and advocacy efforts must spur investment at the national, state, and local levels to implement robust community-based screening programs. Importantly, the 2021 USPSTF lung cancer screening guidelines are estimated to increase eligibility for an additional 6.5 million Americans at high risk, with the greatest increases in eligibility occurring among women and racial minorities. This increase in eligibility presents an opportunity to engage communities that have historically been burdened with the highest rates of lung cancer incidence and mortality and to eliminate disparities in the early detection of lung cancer.

The findings of this study show the impact of lung cancer screening in the US on population-level lung cancer stage shifts and survival. However, given the low adoption of screening previously reported, considering whether low adoption of screening could result in a meaningful stage shift and improvement in survival is important. Previous studies have estimated that between 3.9% and 4.5% of eligible people underwent screening in 2015. If 3.9-4.5% of the 8.0 million eligible people in the US underwent screening each year from 2014 to 2018, an estimated 312,000-360,000 people would have undergone screening each year. If 1% of these people were then diagnosed as having lung cancer, of which 60% were diagnosed at stage I (based on estimates from NLST and NELSON trials), approximately 1600-2160 people would be diagnosed as having stage I disease that was screen detected each year. This increase in the number of NSCLCs detected at stage I would result in a 0.49-0.66% absolute increase in the percentage of stage I disease identified per year, which would translate into a 1.6-2.2% annual percentage increase in the rate of stage I disease identified each year. These percentage estimates, which are theoretically calculated, are within the range of what we observed in our study results. Although other factors may be contributing, the findings of our study suggest that the accelerated increase in the percentage of stage I disease identified beginning in 2014 is likely explained by the start of low-dose computed tomography screening in the US.

Limitations of study
This study has several limitations. Firstly, the NCDB and the SEER Program database do not include data on whether cancers were detected incidentally or via low-dose computed tomography screening. Therefore, we cannot be certain that the accelerated increase in the rate of stage I disease identified beginning in 2014 is a result of the introduction of lung cancer screening in the US. However, we did several analyses to try to exclude the possibility that increases in the rate of stage I disease diagnosed could be explained by factors other than the introduction of lung cancer screening. Secondly, the NCDB and SEER Program database do not include data on patients' smoking history. Therefore, we were able to evaluate only patients potentially eligible for screening on the basis of age. Thirdly, given that this is an observational pre-intervention and post-intervention study, the possibility of confounding exists. However, we tried to reduce the possibility of confounding through the use of multivariable modeling. Fourthly, although patients diagnosed as having lung cancer that was not detected by screening would be the optimal control group, these data were not available and we used patients aged 45-54 who were ineligible for lung cancer screening as a control group instead. However, using patients ages 45-54 as a control group has limitations, given that they have a lower risk of lung cancer and likely have different risk factors for lung cancer compared with older people. Fifthly, in our analysis of low screening states and high screening states, the differences in the screening rate between certain states were small. Sixthly, we obtained the estimates of the number of cases of lung cancer diagnosed in the US by multiplying data from the SEER Program database by 3.57. Additionally, we obtained the estimate of the number of deaths averted due to a shift toward earlier stages identified by multiplying data from the NCDB by 1.54. A limitation of these estimations is that they assume that the samples in the SEER Program database and NCDB are unbiased and representative, which we cannot be certain is true.

Conclusions and policy implications
Although the adoption of lung cancer screening has been slow and screening rates have remained low nationally, the findings of this study indicate the beneficial effect that even a small amount of screening can have on lung cancer stage shifts and survival at the population level. Furthermore, these findings suggest that the widespread adoption of low-dose computed tomography screening should be considered an important strategy for reducing mortality from lung cancer in the US. Efforts to increase utilization of screening should be prioritized to ensure equitable access to screening and to reduce disparities in the stage of lung cancer diagnosed and in survival among different patient populations with lung cancer.

The National Cancer Database (NCDB) data used in the study are derived from a de-identified NCDB file. The American College of Surgeons and the Commission on Cancer have not verified and are not responsible for the analytic or statistical methodology used or the conclusions drawn from these data by the investigators.

Contributors: ALP, ALR, MKV, SAS, HAG, DCC, FIF, and CJY conceived and designed the study, acquired data, interpreted the results, drafted the manuscript, and reviewed or revised the manuscript for important intellectual content. ALP, ALR, and CJY acquired data and did the statistical analyses. All authors interpreted the results and contributed to the drafting and review of the manuscript. All authors approved the final draft. CJY is the guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: None.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/disclosure-of-interest/ and declare no support from any organization for the submitted work, no financial relationships with any organizations that might have an interest in the submitted work in the previous three years, ALP, ALR, and CJY have leadership roles in the American Lung Cancer Screening Initiative.
Ethical approval: This retrospective analysis was approved by the Massachusetts General Hospital Institutional Review Board (IRB #: 2020P004110).

Data sharing: Data used in this study included the National Cancer Database 2018 non-small cell lung cancer file and Surveillance Epidemiology End Results Program database. Data from the National Cancer Database are available through the American College of Surgeons to institutions participating in the National Cancer Database. Data from the Surveillance Epidemiology End Results program database are available at https://seer.cancer.gov/seerstat/download/. Code used for the analyses in this study is available on reasonable request to the corresponding author.

The senior author (the manuscript’s guarantor) affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Dissemination to participants and related patient and public communities: We plan to distribute the work to primary care providers, other clinicians, and lung cancer advocacy groups. In addition, we plan to disseminate the results through social media and as a plain language summary on the American Lung Cancer Screening Initiative website (www.alcs.org).

Provenance and peer review: Not commissioned; externally peer reviewed.

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.


Web appendix: Supplementary materials