Research Square

Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.

Lung Cancer Incidence and Mortality in a Nationwide 45-year Prospective Cohort of Persons with Opioid Use Disorder

Xinyu Wang
Rutgers New Jersey Medical School
Erin Esposito
Rutgers New Jersey Medical School
James J. Goedert
Rutgers New Jersey Medical School
Qiao Zhang
Rutgers New Jersey Medical School
Erin Oxnam
Rutgers New Jersey Medical School
Bridget R. Vaydovsky
Rutgers New Jersey Medical School
Shihui Peng
Rutgers New Jersey Medical School
Andrew W. Bergen
Rutgers New Jersey Medical School
Daniel M. Rosenblum
Rutgers New Jersey Medical School
Stanley H. Weiss
weiss@rutgers.edu

Rutgers New Jersey Medical School

Research Article

Keywords: Opioid use disorder, Lung cancer, Cigarette smoking, Lung cancer screening, Survival analysis, HIV

Posted Date: October 22nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11157942/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Purpose

To examine lung cancer (LC) incidence and mortality in persons with opioid use disorder (OUD) in a national prospective longitudinal cohort.

Methods

Participants (N = 9895) were recruited from US medication-assisted treatment clinics. LC cases and deaths were ascertained through state cancer registries and the National Death Index-Plus. Covariates came from questionnaires and medical records. We calculated standardized incidence ratios (SIRs) and mortality ratios (SMRs) based on general population statistics; survival was estimated using Kaplan-Meier methods.

Results

There were 245 LC cases (56.7% distant stage at diagnosis) and 186 LC deaths, double the expected rates (SIR, 2.03, 95% CI: 1.78–2.30; SMR, 2.19, 95% CI: 1.88–2.52). Risks were similar among non-Hispanic Whites (N = 3227) and non-Hispanic Blacks (N = 3943), higher than Hispanics. SIR and SMR showed a dose-response trend with increasing smoking frequency. Among tobacco smokers with < 20 pack-years, the increased risk was similar to other pack-year groups. Risks were higher with HIV (N = 2079, SIR = 3.31, 95% CI: 2.43–4.40), for both prevalent and incident HIV. Median LC-specific survival was 76 months, 26 months, and 5 months, with local, regional, and distant stage at diagnosis, respectively. LC was frequently diagnosed at late stage.

Conclusion

LC incidence and mortality rates were two-fold elevated. Elevated risks with < 20 pack-years of smoking suggest current LC screening criteria may be too stringent for this population. Integrating LC and tobacco-cessation efforts into drug treatment programs, targeting both providers and patients, may reduce LC morbidity and mortality. Further research with larger, diverse cohorts is needed to confirm these findings.

Introduction

Over 2.1 million people in the United States are affected with Opioid Use Disorder (OUD) [1]. Individuals with OUD experience an elevated risk of adverse health outcomes, such as cardiovascular diseases, cancers, and infectious diseases [2;3]. In particular, people with OUD have an elevated risk for lung diseases, such as pneumothorax, pulmonary edema, and bronchospasms [4]. Among lung-related outcomes, LC is a major cause of morbidity and mortality but has not been examined closely in people with OUD.

LC remains the leading cause of cancer deaths in both males and females, domestically and globally [5]. Prognosis for LC largely depends on the stage of diagnosis, where an earlier stage at diagnosis is associated with an increased likelihood of survival. In contrast, patients diagnosed at an advanced stage face limited treatment options [6]. Based on the American Cancer Society's statistics in 2025, 5-year survival for LC diagnosed at an early stage was 65%, compared to only 10% for those diagnosed at an advanced stage [7]. This underscores the importance of early detection through LC screening to reduce mortality. Notably, early screening with low-dose computed tomography (LDCT) has been shown to reduce LC mortality by approximately 20% [8]. Despite this, less than 20% of those eligible were screened in 2022 [7]. According to the American Lung Association, at initial diagnosis 28% of LCs are localized stage, 41% are regional stage, and 43% are distant stage, with the remaining 8% unstaged [7].

Among patients in treatment for substance addiction, including OUD, more than 80% smoke tobacco compared with only 19.8% among U.S. adults [9;10]. Given that smoking accounts for approximately 85% of LC cases [7] and is disproportionately prevalent among individuals with OUD, there is strong rationale to suspect an elevated LC burden in this population.

To address this gap, we examined LC incidence, mortality, and stage of diagnosis among participants with OUD in a 45-year longitudinal cohort study established and followed by Professor Stanley H. Weiss, MD. We hypothesized that persons with OUD may

have higher LC incidence and mortality rates compared to the general population due to the high prevalence of tobacco use. Based on other work by us and studies by others, including recent focus groups and anonymous surveys of OUD patients by us (manuscript under review: DM Rosenblum, M Dios, E Oxnam, SH Weiss. Cancer screening and education of persons in medication-assisted treatment for opioid use disorder), we expected that LC would be diagnosed at more advanced stages in part related to reduced access to early screening and cancer-specific education [11].

Methods

Cohort Enrollment

In the 1980s, Dr. Weiss, in collaboration with the National Institute on Drug Abuse (NIDA), the National Cancer Institute (NCI), and the New Jersey Department of Health (NJDOH), enrolled 10,223 patients from medication-assisted treatment programs for OUD from across the United States. The study's primary intent was to track the long-term health outcomes of this population over time. These participants completed questionnaires and provided broad consent for the indefinite follow-up of their health outcomes, including access to medical records, except for enrollees from Miami and Chicago for whom follow-up was restricted due to local regulations. Here, we focus on 9,895 unique participants with OUD and sufficient demographic information for follow-up and questionnaire-based analyses.

Data collection

Mortality outcomes were collected via National Death Index Plus (NDI-Plus) matching in 2023 and 2025, with preliminary data through December 31, 2024. Potential death matches were manually adjudicated to ensure specificity and sensitivity. Incident LC cases, including diagnosis date, stage, histology, etc., were collected via 19 state-based cancer registries, facilitated by the Virtual Pooled Registry–Cancer Linkage System (VPR-CLS) [12;13]. These overlapped with the ones identified through death records.

Statistical Analysis

Exposure

Demographic information, drug use, alcohol use, tobacco use, and other information were self-reported. For participants enrolled multiple times, we use baseline data from their first enrollment. Sex information was compared among all of our data sources. Human immunodeficiency virus (HIV) status was determined from the questionnaire, lab results, medical records, death certificate data, and results from state-based AIDS registries. Smoking frequency was used as a measure of smoking intensity. Cigarette pack-years were calculated for the subset of participants with available information. Differences in demographic characteristics, HIV status, and substance use were assessed with Fisher's exact tests for categorical variables and t-tests or Wilcoxon rank-sum tests for continuous variables. For categorical variables with more than two unordered groups, pairwise Fisher's exact tests were performed. Jonckheere-Terpstra tests were used to assess trends across ordered categories. Statistical significance was defined as $p < 0.05$.

Outcomes

The primary outcomes were LC diagnosis and mortality. Mortality was identified in two ways, with LC as (1) only the underlying cause of death, or (2) either the underlying or a contributing cause. The first definition was used for comparison with the NCI's Surveillance, Epidemiology, and End Results (SEER) Program, and the second for all other mortality analyses.

Incident diagnoses were identified through cancer registries and supplemented by death certificate data using primary site cause of death codes from the International Classification of Diseases, 9th Revision (ICD-9) [14;15] — 162, 235.7, 239.1 — and 10th Revision (ICD-10) [16;17] — C34, C39, D38.1. Histologic subtype was classified using the International Classification of Diseases for Oncology, 3rd Edition (ICD-O-3) Histology codes [16;18;19]. The corresponding stage information was based on the NCI's SEER Summary Stage at initial diagnosis. Un-staged cases according to the cancer registry were evaluated based on the proximity of the diagnosis date and the date of death due to LC. For individuals for whom we obtained LC information only from death certificates and the NDI, the

stage was considered missing. Individuals diagnosed with LC prior to enrollment, those diagnosed with in situ stage disease, or those with histology inconsistent with LC were excluded from the analysis.

Standardized Incidence Ratio (SIR) and Standardized Mortality Ratio (SMR) calculations

SIR and SMR were calculated as the ratios of observed to expected cases and deaths, respectively, with 95% confidence intervals calculated using Byar's approximation to the exact Poisson distribution [20]. General population incidence rates (per 100,000) were provided by the New Jersey State Cancer Registry (NJSCR) with confidential, granular tabular data stratified by race and Hispanic ethnicity, sex, 5-year age groups, and calendar year. Per our special request, NJSCR included rates when cases were less than or equal to five. NJSCR began tabulating information specifically on Hispanics only beginning in 1990. For Hispanic participants in the years 1985–1989, we used each year's expected rate for White participants, in accordance with then current NJSCR practices. Cancer incidence data with a comparable level of granularity is not available outside NJ, so we applied the expected number of LC cases calculated using NJ statistics to all study participants regardless of enrollment location. For both incidence and mortality calculations, where a rate within a race/ethnicity-sex group for a given combination of age group and calendar year was not available, we used the rate from the nearest non-missing combination of age group and calendar year.

NJ LC mortality rates for each race/ethnicity, sex, 5-year age group and calendar decade combination were obtained through NCI's SEER Program via CDC Wonder [21]. To align with the SEER mortality definition, SMR analyses were restricted to deaths with LC listed as the underlying cause. LC-caused deaths were identified from death certificate data using the ICD-9 and ICD-10 codes described above.

Expected cancer incidence and mortality within any subset of participants were computed by summing the individuals' cancer incidence and mortality probabilities as determined by the incidence and mortality rates of their respective demographic groups. The most recent NJ mortality rates available was 2022, so 2022 incidence and mortality rates were applied to 2023 and 2024.

Stage Distribution

LC stage was categorized as localized, regional, distant, or unstaged among the cases with available staging information. The stage distribution in our cohort was compared with the SEER stage-at-diagnosis distribution for 2014–2023. SEER-reported stage percentages were converted proportionately to counts summing to 100,000, and a Jonckheere-Terpstra test was used to assess differences in the distribution of stage between our cohort and SEER. For cases with missing or uncertain stage, available diagnosis dates, death dates, medical records, autopsy records and causes of death were reviewed to determine stage at diagnosis.

Cases diagnosed at or near the time of death from LC were categorized as distant stage when the clinical course was considered consistent with advanced disease.

Kaplan-Meier Survival Analysis

Analyses of survival with LC were restricted to cases diagnosed on or prior to the date of death, including six participants whose LC diagnosis date was the same as their date of death. Analyses stratified by stage were further restricted to cases with available stage information. Survival time was calculated from the date of diagnosis to the date of death, else censored as alive at the earlier of the participant's last known follow-up date, or the last date of registry matching, December 31, 2024. Separate survival analyses were conducted for all-cause and LC-specific mortality. For LC-specific survival, the primary analysis defined an event as death with LC listed as either the underlying or a contributing cause. A secondary analysis restricted events to deaths with LC listed as the underlying cause to align with the SEER mortality definition. Survival outcomes were analyzed stratified by race, sex, and cancer stage. Kaplan–Meier survival curves were constructed, and group differences were evaluated using the log-rank and Wilcoxon tests. A two-tailed p-value of < 0.05 was considered statistically significant.

All statistical analyses were performed using SAS software, Version 9.4 (SAS Institute Inc., Cary, NC).

Cigarette Smoking Frequency by LC Histology

Cigarette smoking frequency was assessed using the self-reported average number of cigarettes smoked per day. Smoking frequency distributions were described overall and separately by LC type, including non-small cell LC (NSCLC), small cell LC (SCLC), and cancers with unknown or missing histologic classification. Among NSCLC cases, smoking frequency was further compared across adenocarcinoma, squamous cell carcinoma, and other NSCLC histologic subtypes. Continuous smoking frequency was summarized using means, standard deviations, medians, and interquartile ranges. Differences between histologic groups were evaluated using two-sample t-tests and Wilcoxon rank-sum tests. A two-tailed p-value of < 0.05 was considered statistically significant.

Evaluation of Current Screening Criteria

Lung and bronchus cancer incidence and mortality were further examined according to factors relevant to current low-dose computed tomography (LDCT) screening eligibility. SIRs and SMRs were estimated after stratification by age at enrollment (18–39, 40–49, and ≥ 50 years) and cumulative cigarette smoking exposure (< 20, 20–29, and ≥ 30 pack-years). Because age at enrollment does not reflect age at smoking initiation, smoking frequency (categorized as none/past smoking, < 1 pack/day, and ≥ 1 pack/day) was additionally examined by age group (18–39, 40–49, and ≥ 50 years). Pack-year information was available for only 13.5% of participants, while cigarette smoking frequency was available for 99.6% of participants. Pearson and Spearman correlation coefficients were calculated between cigarette smoking frequency and cumulative smoking exposure measured by pack-years to assess the appropriateness of frequency as a surrogate for cumulative exposure.

Results

Demographic Characteristics

Table 1 lists the demographic characteristics of all OUD participants. Among all 9895 participants, the median age was 35, and participants were followed for a median of 27.4 years. Males made up 64.4% of the cohort and females made up 35.6% of the cohort. Overall, 32.6% of participants self-identified as non-Hispanic (NH) White, 39.8% of participants self-identified as NH Black, and 26.9% of participants self-identified as Hispanic. During follow-up, 245 participants, 70.2% male, were diagnosed with incident LC at a median age of 59 years. The proportion diagnosed with LC did not differ significantly between NH White and NH Black participants (2.73% vs. 3.14%; Fisher's exact p = 0.33), whereas both NH White and NH Black participants had significantly higher proportions of LC than Hispanic participants (1.2%; both p < 0.001) (Table 1). Among participants diagnosed with LC, 96.3% have died, compared to 57.2% of those without LC (Table 1).

In the overall cohort, 90.7% of participants had smoked over 100 cigarettes in their lifetime and 89.5% of participants smoked at the time of study enrollment (Table 1). Baseline smoking prevalence was significantly higher among those who later developed LC compared with those who did not develop LC (97.6% vs 90.5%, Fisher's exact p = 0.0001; Table 1). Smoking frequency differed significantly between participants who developed LC and those who did not (trend p < 0.0001), with a higher proportion of LC cases reporting ≥ 1.5 packs/day (38.0% vs. 24.0%).

Among the 1006 participants who enrolled more than once, reported smoking changed very little. Between first and last enrollments [median interval, 1.23 years (Interquartile range (0.74, 2.06) years], median change in cigarettes per day (CPD) was 0; change was − 5 to + 5 CPD among 68.3%, and − 10 to + 10 CPD among 83.0%. There was no overall trend toward a decrease or an increase (p = 0.63, two-tailed Wilcoxon signed-rank test). Furthermore, among the 1323 participants with available information, pack-years were strongly correlated with cigarette smoking frequency (Pearson r = 0.79; Spearman r = 0.81; both p < 0.0001), thus further supporting use of baseline cigarette smoking frequency as a measure of overall cigarette exposure.

A total of 41.6% of participants had less than a high school education (including nine who reported no formal education). At the time of study enrollment, 97% had used heroin and 84.2% had used cocaine (Table 1). Among participants with LC, 92.7% had used intravenous (IV) heroin during the previous 5 years, with a similar proportion among participants without LC (92.3%; Fisher's exact p = 0.06). A similar trend was observed for IV heroin use frequency during the previous 5 years.

Stage at diagnosis

Of the 245 LC cases identified, state-based cancer registries provided stage at diagnosis for 143 cases (58.4%). Among these cases, 15.4% were localized, 26.6% regional, 49.0% distant, and 9.1% unstaged, compared with 26.3%, 20.3%, 46.5%, and 6.9%, respectively,

in SEER data from 2014–2023 (Table 2). The stage distributions differed significantly, with our cohort showing a shift toward more advanced stage at diagnosis compared with SEER (one-tailed Jonckheere-Terpstra trend test, $p = 0.019$). Stage was determined for 11 additional cases (10 distant, 1 regional) by reviewing diagnosis dates, death dates, medical records, and causes of death. Among the ten cases we categorized as distant stage at diagnosis, seven were diagnosed at the time of death from LC, two were first diagnosed 54 and 58 days before death from LC had a clinical course of disease aligned with distant stage at diagnosis, and one was identified through autopsy, which confirmed metastatic LC with spread to the brain. The case we recategorized from unstaged to regional stage at diagnosis had AJCC stage data T4N1M0 reported in the cancer registry despite missing SEER stage data.

Stage at diagnosis and histological subtypes of LC were compared for the 121 cases that had both histology data and stage data including those that were recategorized. Sixteen individuals had small cell LC (SCLC), all at distant or regional stage, and the remaining 105 had non-small cell LC (NSCLC), including 18 diagnosed at localized stage.

Stage at diagnosis did not differ by sex, but it was inversely related to heroin use and frequency within the past 5 years (both $p < 0.05$). Frequent heroin use (≥ 1/week) was reported by 95.5% of localized, 92.3% of regional, and 78.8% of distant stage cases (Jonckheere-Terpstra $p = 0.024$; Table 3).

The proportion of participants known to be dead increased with more advanced stage at diagnosis (trend $p < 0.0001$), with 98.8% of those diagnosed with distant-stage LC known to be deceased compared with 84.6% of those diagnosed with regional LC (Table 3).

SIR & SMR Analyses

As shown in Table 4, considering all 245 LC cases in our cohort versus the general population, there was a significantly elevated standardized incidence ratio (SIR) of 2.03 (95% CI: 1.78–2.30) and standardized mortality ratio (SMR) of 2.19 (95% CI: 1.88–2.52). SIR and SMR were similarly elevated across age groups (18–39, 40–49, and 50 + years). By sex, SIR was slightly (1.2-fold) higher among females (2.32, 95% CI: 1.82–2.91) than males (1.93, 95% CI: 1.65–2.24). Similarly, SMR was 1.4-fold higher among females (2.79, 95% CI: 2.09–3.63) than males (2.01, 95% CI: 1.68–2.38). By race/ethnicity, LC incidence and mortality were more than 2-fold higher among NH Whites and NH Blacks than among Hispanics (Table 4). Participants with less than a high school education had approximately 1.5-fold higher SIR and SMR than those with a higher education (Table 4). Participants with HIV (combining prevalent and incident HIV) had an elevated SIR (SIR = 3.31, 95% CI: 2.43–4.40), much higher than participants who remained HIV negative throughout study duration (SIR = 1.86, 95% CI: 1.61–2.14).

Stratification by smoking frequency revealed a dose–response trend in SIR and SMR. Participants who were not current smokers had a low SIR (0.44, 95% CI: 0.14–1.02) while participants who smoked < 1 pack/day, between 1 to 1.5 packs/day, and more than 1.5 packs/day had stepwise increases in SIR (1.59, 95% CI: 1.17–2.10; 2.03, 95% CI: 1.65–2.48; 3.35, 95% CI: 2.70–4.10), respectively. A dose-response trend with smoking frequency was also seen in SMR. Notably, LC incidence was significantly elevated in both younger age-at-enrollment groups (18–39 and 40–49 years). LC incidence and mortality were significantly elevated even in the subset of 808 participants with < 20 pack-years of cumulative smoking (SIR, 2.14; 95% CI, 1.27–3.39; SMR, 2.62; 95% CI, 1.43–4.40) (Table 4).

Cigarette Smoking Frequency by Age at Diagnosis and LC Histology

Approximately 20% of all cases were diagnosed before age 50. Smoking intensity was high among cases diagnosed before the current screening age threshold: smoking ≥ 1 pack/day was reported by 76.7% and 83.3% of cases diagnosed at ages 40–49 and 18–39, respectively. Average cigarette consumption was similar across LC histologic groups. Among cases, mean cigarette consumption was 24.71 CPD with NSCLC and 25.56 CPD with SCLC, with a median of 20 CPD in both groups. Cigarette consumption did not differ significantly between NSCLC and SCLC (t-test p-value = 0.80; Wilcoxon two-tailed $p = 0.61$). With NSCLC adenocarcinoma (N = 59), mean consumption was 23.56 CPD; with squamous cell carcinoma (n = 28), 26.07 CPD; and with other NSCLC subtypes (n = 21), 26.14 CPD. No statistically significant differences were observed between adenocarcinoma and squamous cell carcinoma (t-test p-value = 0.40; Wilcoxon p-value = 0.32).

Survival Analyses

Participants with LC as an underlying or contributing cause of death, including those with an unspecified cancer-related death whose only known cancer diagnosis was LC, the median survival time was 11 months (Fig. 1). Stratified by stage at diagnosis, the median survival among participants with local, regional, and distant stage LC at diagnosis was 76 months, 26 months, and 5 months, respectively (Fig. 2). These estimates differed minimally from those expected from SEER (Supplemental Figs. 1 and 2). Median all-

cause survival for participants with LC was 9 months (Supplemental Fig. 3). All-cause survival differed significantly by stage at diagnosis (Supplemental Fig. 4; log-rank $p < 0.0001$). Among the participants excluded from this analysis, the most common cause of death was primary cancer of a different specified site, followed by fatal overdose.

Discussion

Overall risk, prevention, and screening

In conclusion, we observed a 2-fold higher incidence rate and a 2.2-fold higher mortality rate of LC among persons living with OUD compared to the general population. The increased incidence rate is likely attributable to the high prevalence of smoking within our cohort (90%), which is further supported by the dose response relationship between smoking frequency and SIR. However, other factors associated with OUD may also contribute. In U.S. cohorts with or at high risk of HIV, LC incidence increased with older age, was 5-fold higher with ≥ 10 smoking pack-years, and was 2-fold higher with severe HIV immunodeficiency [22]. Injection drug use was associated with approximately 2-fold higher incidence in univariate but not multivariable analysis [22]. The elevated mortality rate observed in our cohort may be related to the high proportion of cases diagnosed at a distant stage (56.7%). Surprisingly, stage at diagnosis was inversely related to injection of heroin, an observation without a ready explanation. In addition to our findings, Sheikh et al. reported that opium smoking was associated with an increased risk of LC in northeast Iran, with a median follow-up of 10 years, even among non-smokers of tobacco [23]. Another study examined the association between the duration of illicit drug use and various health conditions, including LC, but found no correlation between LC and drugs other than cannabis [24]. However, that study was restricted to individuals aged 35 to 49 years, who may still have been within the latency period for LC development. An Australian death index linkage study found higher LC mortality among opioid substitution therapy registrants than in the general population (SMR 3.6, 95% CI 2.8–4.6) [25]. Similarly, a Norwegian national study reported increased LC incidence (SIR 3.5, 95% CI 2.8–4.3) and mortality (SMR 3.9, 95% CI 3.0–5.0) among individuals with OUD [26]. These estimates were higher than those observed in our cohort, although all studies showed an elevated LC burden among persons with OUD. The causal relationship between opioids and LC, if any, remains uncertain. For example, a Mendelian randomization analysis found no causal link between OUD and LC [27] in the UK BioBank and FinnGen populations, though the UK BioBank cohort has lower rates of poverty and smoking than the general UK population [27;28].

Although we found elevated risks, the differences between the SIR and SMR were minimal, which could suggest a potential underutilization of LC screening. The high proportion of individuals diagnosed at an advanced stage further supports this possibility. This raises the question of whether OUD participants are being omitted from or lack access to LC screening altogether. In a study of 911 high-risk participants who underwent LDCT screening for LC, 14 of them were found to have LC, of whom 13 were diagnosed with stage one cancer [8]. LDCT is clearly an effective mechanism to diagnose LC at an early stage while there are more treatment options available. OUD treatment centers and clinicians should together work to ensure OUD patients are being screened regularly to increase the chances of catching LC while it is treatable.

Many people treated for OUD also use multiple drugs besides opioids. It is estimated that between 70 and 90% of patients in medication-assisted treatment programs smoke tobacco [29]. Around the year 2000, the corresponding author of this paper demonstrated increased LC cases among drug users and shared the findings with the NJDOH [30;31]. Following this, the treatment centers began routinely recommending tobacco cessation referrals. However, as illustrated by the null change in smoking frequency that we found here, only 40% of substance abuse treatment centers offer smoking cessation counseling and only 17% offer medications to assist with smoking cessation [29]. The issue of smoking and its associated LC risk has persisted for decades.

The criteria for LC screening were expanded in 2021, lowering the recommended screening age from 55 years to 50 years, and the cumulative smoking exposure threshold from 30 to 20 pack-years [32]. Low-dose CT screening has been shown to reduce LC mortality by approximately 20% compared with chest radiography [8]. However, in our study, 18 (54.55%) of 33 LC cases with available pack-year information had < 20 pack-years of cumulative smoking exposure; thus, these cases fell below the current smoking exposure threshold for LDCT screening. Additionally, this group had significantly elevated LC incidence compared with the general population, with an SIR similar to the 20–29 pack-year group. LC mortality was also significantly elevated among participants with < 20 pack-years, whereas SMR estimates for the higher pack-year groups did not reach statistical significance, likely due to small numbers. These findings suggest that the current 20-pack-year threshold for LDCT screening may not fully capture elevated LC risk among persons with OUD.

In addition, the median age of LC diagnosis in our cohort (59 years) was younger than the average age of LC diagnosis in the general population (70 years) [33]. Although these findings do not establish that the screening age should be lowered, they suggest that current screening criteria for age may not identify all high-risk people with OUD and support further research on whether earlier or risk-based screening criteria for this population.

Other drugs and suicide

Persons using cocaine do not appear to have an increased incidence of LC [24]. Cocaine use was not associated with LC incidence in our study. In our cohort, individuals with OUD frequently engaged in multiple drug use. The combined effects of multiple illicit drugs on LC risk remain unclear, and it is possible that these substances exert additive effects that increase risk.

Alcohol use was not found to be associated with an increased risk of LC diagnosis in our cohort. Previous studies regarding this possible association have mixed results but often report that any elevated LC risk is dependent on the quantity of alcohol consumption and the type of alcohol consumed. A pooled analysis of 22 studies on LC among non-smokers found low to moderate consumption of wine or liquor to be inversely associated with LC risk [34]. However, a systematic review of 24 studies found that beer and liquor consumption are associated with an increased risk of LC [35].

We had 4 study participants with LC who died of a fatal overdose. One subject died of an overdose on the date of LC diagnosis. One study reported that cancer patients had 4.4 times the chance of committing suicide compared to the general population [36]. This highlights an important clinical implication: the need to remain sensitive and take additional care while delivering news of a cancer diagnosis to all patients, especially patients with a psychiatric condition such as OUD.

HIV-infection

We did find elevated SIR and SMR among HIV-infected individuals (Table 4). This finding aligns with previous research. For example, Kirk et al. reported a hazard ratio of 3.6 (95% CI: 1.6–7.9) after adjusting for age, sex, smoking status, and calendar period [37]. Similarly, Shiels et al found that among drug users, HIV-infected individuals had approximately twice the risk of LC compared to HIV-negative individuals [38]. In a study of U.S. high-risk cohorts, LC incidence was at least 2-fold higher with severe HIV immunodeficiency [22]. Currently neither OUD nor HIV status are considered in LC screening recommendations This omission may lead primary care providers to overlook the elevated risk in these groups, potentially resulting in failure to recommend early screening and contributing to late-stage diagnoses, limited treatment options, and poor outcomes.

Sex, race/ethnicity, and education

In the general population, male LC incidence has been higher than female incidence [39–41], but rates of specific LCs such as adenocarcinoma are currently rising among younger females and females who do not smoke. No LC screening guidelines currently exist for these populations. Our study suggests that women with OUD have an incidence of LC significantly increased compared to the general population, and perhaps higher than men with OUD, similar to recent findings in the general population [42]. A sex difference in LC risk was not found in U.S. cohorts with or at high risk of HIV [22]. Whether sex disparities exist at the stage of diagnosis remains controversial. Araghi et al. found that males were more likely to be diagnosed at an advanced stage [43], while the current study and others found no sex differences in stage at diagnosis [44;45]. Tolwin et al. observed that smoking females were more likely to be diagnosed at a later stage [46], suggesting that some individuals who smoke may delay seeking medical attention despite experiencing alarming symptoms [47;48].

Many studies have demonstrated that lower socioeconomic status (SES), including education as a key component, is strongly associated with higher LC incidence and poorer survival [49–51]. Consistent with these findings, our cohort analysis showed that individuals with less than a high school education had an SIR of 2.68 and an SMR of 2.74 compared with the general population. Thus, education and other components of SES should be considered in tobacco and LC prevention screening efforts.

Limitations

This study has several limitations. First, the number of LC cases within our cohort limited statistical power to evaluate associations between specific frequencies of drug types (e.g., opioids, cocaine) and LC risk. In addition, the available social determinants of health variables were limited in scope, constraining our ability to examine the role of factors such as insurance, income, housing stability, and occupational exposures.

Conclusion & Future Direction

Our findings indicate that individuals with OUD experience substantially higher LC incidence and mortality rates compared with the general population. Together with the high prevalence and intensity of cigarette smoking observed in this cohort, these results suggest that lowering the age for initial LC screening for people with OUD may warrant consideration. While our data cannot establish causality, they highlight the need for further evaluation of whether existing LC screening criteria adequately capture this potentially high-risk group.

People with opioid use disorder attending medication-assisted treatment programs present a unique opportunity for repeated access to a high-risk population. The absence of a significant change in smoking frequency in sequential enrollments highlights the unfulfilled potential value of smoking cessation efforts among persons in these programs. Implanting LC education and smoking cessation resources into drug use treatment programs, as well as increasing healthcare provider awareness of referrals for LC screening for drug users, may help reduce the LC burden in this population. Broader collaboration between treatment center and clinicians is needed, and current policies should be revisited to ensure this high-risk group is not being overlooked as to LC screening.

Continued investigation with larger, geographically diverse cohorts and more detailed socioeconomic data will be crucial to further define risk and inform future screening recommendations. As a model, this project demonstrated collaboration between academics and state cancer registries and calls attention to an area with opportunities for public health and clinical intervention.

Declarations

Funding

This work was supported by the National Institute on Drug Abuse of the National Institutes of Health under Award Numbers R01DA050495 and R01DA044014, the New Jersey Commission on Cancer Research under Award Number COCR26PPR013, and the Rutgers Global Health Institute through a Rutgers Global Health Seed Grant.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Author Contributions

Conceptualization: Stanley H. Weiss; Methodology: Stanley H. Weiss, Xinyu Wang, Erin Esposito, James J. Goedert, Qiao Zhang, Daniel M. Rosenblum, Andrew W. Bergen; Formal analysis and investigation: Stanley H. Weiss, Xinyu Wang, Erin Esposito, Qiao Zhang, Daniel M. Rosenblum, Erin Oxnam, Bridget R. Vaydovsky, Shihui Peng; Writing – original draft preparation: Qiao Zhang; Writing – review and editing: Stanley H. Weiss, Xinyu Wang, Erin Esposito, James J. Goedert, Andrew W. Bergen; Funding acquisition: Stanley H. Weiss, Daniel M. Rosenblum; Resources: Stanley H. Weiss, Daniel M. Rosenblum; Supervision: Stanley H. Weiss, James J. Goedert, Daniel M. Rosenblum, Andrew W. Bergen

Data Availability

The data that support the findings of this study are not openly available due to reasons of sensitivity and confidentiality, including but not limited to the facts that participants' consent to participate included an assurance of maintenance of confidentiality of the data and that some data were acquired from U.S. state cancer registries subject to data usage agreements that prohibit further sharing.

Ethics approval

This research was conducted in line with the principles of the Declaration of Helsinki. Approval was granted by the Rutgers Institutional Review Board under protocol Pro20150001314, initially approved on 17 June 2015 and with continuing reviews approved each year since then.

Consent to participate

Informed consent was obtained from all individual participants included in the study.

Acknowledgements:

Additional support was provided by the Rutgers New Jersey Medical School Summer Student Research Program (Newark, NJ) and internal resources. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the New Jersey Commission on Cancer Research, the Rutgers Global Health Institute, or the New Jersey Department of Health.

We thank the participants, who volunteered for this study, and the many treatment programs for their collaboration, as well as the students, staff and collaborators who have assisted us over time with this project, especially Jinu (Brandon) Son, Daniel Ramge, Susan B. German, MPH, and Gerald Harris, PhD.

References

1. Dydyk AM, Jain NK, Gupta M (2020) Opioid Use Disorder. StatPearls Publishing, Treasure Island, FL
2. Blanco C, Volkow ND (2019) Management of opioid use disorder in the USA: present status and future directions. Lancet 393(10182):1760–1772
3. Hser YI, Mooney LJ, Saxon AJ et al (2017) High Mortality Among Patients With Opioid Use Disorder in a Large Healthcare System. J Addict Med 11(4):315–319
4. Radke JB, Owen KP, Sutter ME, Ford JB, Albertson TE (2014) The Effects of Opioids on the Lung. Clin Rev Allergy Immunol 46(1):54–64
5. World Health Organization (2026) Lung cancer. World Health Organization
6. National Comprehensive Cancer Network (2026) NCCN Guidelines for Patients: Early and Locally Advanced Non-Small Cell Lung Cancer. National Comprehensive Cancer Network, Plymouth Meeting, PA, pp 1–82
7. American Lung Association (2025) State of Lung Cancer 2025 Report.55 W. Wacker Drive, Suite 1150, vol 60601. American Lung Association, Chicago, IL
8. The National Lung Screening Trial Research Team (2011) Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening. N Engl J Med 365(5):395–409
9. Guydish J, Passalacqua E, Pagano A et al (2016) An international systematic review of smoking prevalence in addiction treatment. Addiction 111(2):220–230
10. CDC (2024) Tobacco Product Use Among Adults-United States, 2022.Washington, DC. Centers for Disease Control and Prevention (CDC)
11. Dios M, Rosenblum DM, Weiss SH (2026) Examining the Deficiencies in Education and Screening for Cancer Among Men and Women. Medication-Assisted Drug Treatment Programs in Five New Jersey Counties - Final Project Report. Rutgers Global Health Institute, Newark, NJ, pp 1–12
12. Deapen D (2016) Centralized access to cancer registries. Ann Epidemiol 26(11):754–756
13. Deapen D, Clerkin C, Howe W et al (2024) Virtual Pooled Registry-Cancer Linkage System: an improved method for ascertaining cancer diagnoses. J Natl Cancer Inst Monogr 2024(65):191–197
14. World Health Organization (1978) International classification of diseases: [9th] ninth revision, basic tabulation list with alphabetic index. World Health Organization, Geneva
15. Mortality Statistics Branch, National Center for Health Statistics (1998) ICD-9 Volume 1, Modified for US Purposes. Centers for Disease Control and Prevention, Hyattsville, MD
16. World Health Organization (2000) International Classification of Diseases for Oncology, 3rd Edition (ICD-O-3). World Health Organization
17. World Health Organization (2026) International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10 Version:2019)
18. National Cancer Institute Surveillance, Epidemiology, and End Results (SEER) Program (2026) SEER ICD-O-3 Coding Materials. National Cancer Institute
19. North American Association of Central Cancer Registries (NAACCR) (2026) ICD O 3 Coding Updates. North American Association of Central Cancer Registries (NAACCR)

20. Breslow NE, Day NE (1987) Statistical Methods in Cancer Research: Volume II - The Design and Analysis of Cohort Studies, 82 edn. International Agency for Research on Cancer, Lyon, France
21. Centers for Disease Control and Prevention (2026) CDC WONDER. Centers for Disease Control and Prevention
22. Hessol NA, Mart+¡nez-Maza O, Levine AM et al (2015) Lung cancer incidence and survival among HIV-infected and uninfected women and men. AIDS 29(10):1183–1193
23. Warnakulasuriya S, Cronin-Fenton D, Jinot J et al (2020) Carcinogenicity of opium consumption. Lancet Oncol 21(11):1407–1408
24. Han B, Gfroerer JC, Colliver JD (2010) Associations Between Duration of Illicit Drug Use and Health Conditions: Results from the 2005–2007 National Surveys on Drug Use and Health. Ann Epidemiol 20(4):289–297
25. Randall D, Degenhardt L, Vajdic CM et al (2011) Increasing cancer mortality among opioid-dependent persons in Australia: a new public health challenge for a disadvantaged population. Aust N Z J Public Health 35(3):220–225
26. Kostovski E, Hamina A, Hjellvik V, Clausen T, Skurtveit S (2024) Increased cancer incidence and mortality among people with opioid use-related disorders: A nation-wide cohort study. Addiction 119(9):1575–1584
27. Su X, Mo X, Kan J, Yang F, Zhang B, Huang Y (2025) Decoding the genetic links between substance use disorder and cancer vulnerability. Psychopharmacology 242(9):2021–2033
28. Fry A, Littlejohns TJ, Sudlow C et al (2017) Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. Am J Epidemiol 186(9):1026–1034
29. Friedmann PD, Jiang L, Richter KP (2008) Cigarette smoking cessation services in outpatient substance abuse treatment programs in the United States. J Subst Abuse Treat 34(2):165–172
30. Weiss SH, Knox K, Roche L, Abe T (2002) Increased cancer incidence in NJ injection drug users (IDU). (Plenary session D3.).Princeton, New Jersey: p. 38
31. Weiss SH, Karcnik T, Patel S et al (1997) The ten-year cancer and HIV experience of the UMDNJ-University Hospital cancer registry: excess proportions of Hodgkin's lymphoma and lung cancer. J Acquir Immune Defic Syndr Hum Retrovirology 14(4):A18
32. US Preventive Services (2021) Screening for Lung Cancer: US Preventive Services Task Force Recommendation Statement. JAMA 325(10):962–970
33. American Cancer Society (2026) Key Statistics for Lung Cancer. American Cancer Society, Atlanta, Georgia
34. Fehringer G, Brenner DR, Zhang ZF et al (2017) Alcohol and lung cancer risk among never smokers: A pooled analysis from the international lung cancer consortium and the SYNERGY study. Int J Cancer 140(9):1976–1984
35. Bandera EV, Freudenheim JL, Vena JE (2001) Alcohol Consumption and Lung Cancer: A Review of the Epidemiologic Evidence. Cancer Epidemiol Biomarkers Prev 10(8):813–821
36. Zaorsky NG, Zhang Y, Tuanquin L, Bluethmann SM, Park HS, Chinchilli VM (2019) Suicide among cancer patients. Nat Commun 10(1):207
37. Kirk GD, Merlo C, O'Driscoll P et al (2007) HIV Infection Is Associated with an Increased Risk for Lung Cancer, Independent of Smoking. Clin Infect Dis 45(1):103–110
38. Shiels MS, Cole SR, Mehta SH, Kirk GD (2010) Lung cancer incidence and mortality among HIV-infected and HIV-uninfected injection drug users. J Acquir Immune Defic Syndr 55(4):510–515
39. Fu JB, Kau TY, Severson RK, Kalemkerian GP (2005) Lung Cancer in Women: Analysis of the National Surveillance, Epidemiology, and End Results Database. Chest 127(3):768–777
40. Kinoshita FL, Ito Y, Morishima T, Miyashiro I, Nakayama T (2017) Sex differences in lung cancer survival: long-term trends using population-based cancer registry data in Osaka, Japan. Jpn J Clin Oncol 47(9):863–869
41. Tabatabai MA, Kengwoung-Keumo JJ, Oates GR et al (2016) Racial and Gender Disparities in Incidence of Lung and Bronchus Cancer in the United States: A Longitudinal Analysis. PLoS ONE 11(9):e0162949
42. Ragavan MV, Patel MI (2020) Understanding Sex Disparities in Lung Cancer Incidence: are Women more at Risk? Lung Cancer Manag 9(3):LMT34
43. Araghi M, Fidler-Benaoudia M, Arnold M et al (2022) International differences in lung cancer survival by sex, histological type and stage at diagnosis: an ICBP SURVMARK-2 Study. Thorax 77(4):378–390
44. Visbal AL, Williams BA, Nichols FC et al (2004) Gender differences in non-small-cell lung cancer survival: an analysis of 4,618 patients diagnosed between 1997 and 2002. Ann Thorac Surg 78(1):209–215

45. Quillet A, Le Stang N, Meriau N, Isambert N, Defossez G (2024) Socio-demographic inequalities in stage at diagnosis of lung cancer: A French population-based study. Cancer Epidemiol 89:102522
46. Tolwin Y, Gillis R, Peled N (2020) Gender and lung cancer—SEER-based analysis. Ann Epidemiol 46:14–19
47. Friedemann Smith C, Whitaker KL, Winstanley K, Wardle J (2016) Smokers are less likely than non-smokers to seek help for a lung cancer 'alarm' symptom. Thorax 71(7):659–661
48. Stapelfeld C, Dammann C, Maser E (2020) Sex-specificity in lung cancer risk. Int J Cancer 146(9):2376–2382
49. Hovanec J, Kendzia B, Olsson A et al (2025) Socioeconomic Status, Smoking, and Lung Cancer: Mediation and Bias Analysis in the SYNERGY Study. OVID 36(2):245–252
50. Zhou H, Zhang Y, Liu J et al (2019) Education and lung cancer: a Mendelian randomization study. Int J Epidemiol 48(3):743–750
51. Zou X, Wang R, Yang Z et al (2022) Family Socioeconomic Position and Lung Cancer Risk: A Meta-Analysis and a Mendelian Randomization Study. Front Public Health 10(2022): 780538

Tables

Table 1
Demographic and behavioral characteristics of 9,895 enrolled participants with opioid use disorder.

Age at enrollment (median, range)	**Total (n = 9895)**		**Lung cancer cases (n = 245)**		**Participants without known lung cancer (n = 9650)**	
	35	**18–79**	**38**	**22–66**	**34**	**18–79**
Years of follow-up (median, range)	27.4	0.01–44.09	20.5	1.57–40.22	27.84	0.01–44.09
Age at diagnosis (median, range)	n/a	n/a	59	35–87	n/a	n/a
Vital Status (n, Column %)						
Not known dead	4140	41.8%	14	5.7%	4126	42.8%
Dead	5755	58.2%	231	94.3%	5524	57.2%
Sex (n, Column %)						
Male	6371	64.4%	172	70.2%	6199	64.2%
Female	3524	35.6%	73	29.8%	3451	35.8%
Race/Ethnicity (n, Column %)						
NH White	3227	32.6%	88	35.9%	3139	32.5%
NH Black	3943	39.8%	124	50.6%	3819	39.6%
Hispanic	2665	26.9%	32	13.1%	2633	27.3%
Other Race	60	0.6%	1	0.4%	59	0.6%
HIV Status (n, Column %)						
HIV positive (prevalent and incident cases)	2079	21.0%	47	19.2%	2032	21.1%
HIV positive at baseline	1716	17.3%	38	15.5%	1678	17.4%
Incident HIV	363	3.7%	9	3.7%	354	3.7%
HIV negative at baseline without incident HIV diagnosis	7759	78.4%	196	80.0%	7563	78.4%
Unknown	57	0.6%	2	0.8%	55	0.6%
Education Level (n, Column %)						
<High School	4119	41.6%	113	46.1%	4006	41.5%
=High School	3594	36.3%	74	30.2%	3520	36.5%
>High School	2023	20.4%	56	22.9%	1967	20.4%
Not ascertained	159	1.6%	2	0.8%	157	1.6%
Tobacco Use, ever used > 100 in lifetime (n, Column %)						
Yes	8972	90.7%	239	97.6%	8733	90.5%
No	781	7.9%	5	2.0%	776	8.0%
Not ascertained	142	1.4%	1	0.4%	141	1.5%
Tobacco Use, frequency (n, Column %)						
Non-current smoker[a]	820	8.3%	5	2.0%	815	8.4%
Current smoker	8858	89.5%	239	97.6%	8619	89.3%

Age at enrollment (median, range)	**Total (n = 9895)**		**Lung cancer cases (n = 245)**		**Participants without known lung cancer (n = 9650)**	
	35	**18–79**	**38**	**22–66**	**34**	**18–79**
< 1 pack/day	2566	25.9%	49	20.0%	2517	26.1%
1-1.5 pack/day	3883	39.2%	97	39.6%	3786	39.2%
≥ 1.5 pack/day	2409	24.3%	93	38.0%	2316	24.0%
Not ascertained	217	2.2%	1	0.4%	216	2.2%
Alcohol Use Ever (n, Column %)						
Yes	7233	73.1%	180	73.5%	7053	73.1%
No	2512	25.4%	63	25.7%	2449	25.4%
Not ascertained	150	1.5%	2	0.8%	148	1.5%
Opioid Use, frequency (n, Column %)						
≤ 1/day	1254	12.7%	24	9.8%	1230	12.7%
> 1/day, ≤ 2.5/day	1881	19.0%	57	23.3%	1824	18.9%
> 2.5/day, ≤ 4/day	2662	26.9%	75	30.6%	2587	26.8%
> 4/day	3742	37.8%	77	31.4%	3665	38.0%
Not ascertained	356	3.6%	12	4.9%	344	3.6%
Heroin Ever Use (n, Column %)						
Yes	9595	97.0%	233	95.1%	9362	97.0%
No	216	2.2%	10	4.1%	206	2.1%
Not ascertained	84	0.8%	2	0.8%	82	0.8%
Cocaine use in last 5 years (n, Column %)						
Yes	8331	84.2%	202	82.4%	8129	84.2%
No	1475	14.9%	41	16.7%	1434	14.9%
Not ascertained	89	0.9%	2	0.8%	87	0.9%

[a] Participants enrolled in NJ were asked about both their past smoking habits and their current smoking habits. Thirty-nine participants enrolled in NJ reported that they used to smoke in the past but no longer smoke. Participants enrolled in states other than NJ were only asked about their current smoking habits. Because the past smoker data was incompletely ascertained for the overall cohort, all NJ enrolled past smokers are grouped with participants who were not current smokers at the time of study enrollment for analysis.

Table 2
Proportion analysis of cancer stage at diagnosis among participants with stage at diagnosis information directly from the cancer registry compared to SEER data for the country. This analysis does not take our reclassifications into consideration.

	Localized		Regional		Distant		Unstaged	
	number of participants	percent of diagnoses	number of participants	percent of diagnoses	number of participants	percent of diagnoses	number of participants	percent of diagnoses
Distribution by stage at dx	22	15.40%	38	26.60%	70	49.00%	13	9.10%
SEER stage at dx (2014–2023)		26.3%		20.30%		46.50%		6.90%

Jonckheere-Terpstra Test for trend: **one-tailed p value = 0.0191**.

Table 3
Demographic and behavioral characteristics of 141 participants with opioid use disorder (OUD) who had cancer stage information available at diagnosis. As detailed in our results, seven unstaged cases were re-categorized as distant stage and one was categorized as regional stage.

	Total lung cancer cases with stage information		**Local**		**Regional**		**Distant**	
	(n = 141)		**(n = 22, 15.6%)**		**(n = 39, 27.7%)**		**(n = 80, 56.7%)**	
Age at enrollment (median, range)	37	22–66	39.5	27–66	35	22–54	37	23–63
Years of follow-up (median, range)	21.9	1.60-40.22	26.48	7.56–37.76	26.26	2.60-40.22	19.47	1.73–33.68
Age at diagnosis of lung cancer (median, range)	57	35–82	65.5	45–80	56	40–78	57.5	35–82
Vital Status (n, Column %)								
Dead	127	90.1%	15	68.2%	33	84.6%	79	98.8%
Not known dead	14	9.9%	7	31.8%	6	15.4%	1	1.3%
Sex (n, Column %)								
Male	93	66.0%	14	63.6%	27	69.2%	52	65.0%
Female	48	34.0%	8	36.4%	12	30.8%	28	35.0%
Race/Ethnicity (n, Column %)								
NH White	49	34.8%	7	31.8%	11	28.2%	31	38.8%
NH Black	73	51.8%	10	45.5%	22	56.4%	41	51.3%
Hispanic	19	13.5%	5	22.7%	6	15.4%	8	10.0%
HIV Status (n, Column %)								
HIV positive at baseline	28	19.9%	5	22.7%	9	23.1%	14	17.5%
Incident HIV	7	5.0%	1	4.5%	1	2.6%	5	6.3%
HIV negative at baseline without incident HIV diagnosis	104	73.8%	16	72.7%	29	74.4%	59	73.8%
Unknown	2	1.4%	0	0.0%	0	0.0%	2	2.5%
Education Level (n, Column %)								
<High School	69	48.9%	13	59.1%	20	51.3%	36	45.0%
=High School	43	30.5%	6	27.3%	14	35.9%	23	28.8%
>High School	27	19.1%	3	13.6%	5	12.8%	19	23.8%
Not ascertained	2	1.4%	0	0.0%	0	0.0%	2	2.5%
Tobacco Use, ever used > 100 in lifetime(n, Column %)								
Yes	138	97.9%	21	95.5%	38	97.4%	79	98.8%
No	2	1.4%	1	4.5%	1	2.6%	0	0.0%
Not ascertained	1	0.7%	0	0.0%	0	0.0%	1	1.3%
Tobacco Use, frequency (n, Column %)[a]								
Non-current smoker	2	1.4%	1	4.5%	1	2.6%	0	0.0%
Smoker	138	97.9%	21	95.5%	38	97.4%	79	98.8%

	Total lung cancer cases with stage information		Local		Regional		Distant	
	(n = 141)		(n = 22, 15.6%)		(n = 39, 27.7%)		(n = 80, 56.7%)	
< 1 pack/day	25	17.7%	7	31.8%	6	15.4%	12	15.0%
1-1.5 pack/day	56	39.7%	6	27.3%	13	33.3%	37	46.3%
≥ 1.5 pack/day	57	40.4%	8	36.4%	19	48.7%	30	37.5%
Not ascertained	1	0.7%	0	0.0%	0	0.0%	1	1.3%
Alcohol Use (n, Column %)[b]								
Yes	106	75.2%	16	72.7%	34	87.2%	56	70.0%
No	33	23.4%	6	27.3%	5	12.8%	22	27.5%
Not ascertained	2	1.4%	0	0.0%	0	0.0%	2	2.5%
Alcohol Use, frequency (n, Column %)[b]								
Abstain (0 + < 0.2)	54	38.3%	12	54.5%	10	25.6%	32	40.0%
Drinker	85	60.3%	10	45.5%	29	74.4%	46	57.5%
Light	40	28.4%	6	27.3%	15	38.5%	19	23.8%
Moderate	18	12.8%	1	4.5%	4	10.3%	13	16.3%
Heavy	27	19.1%	3	13.6%	10	25.6%	14	17.5%
Missing	2	1.4%	0	0.0%	0	0.0%	2	2.5%
Opioid Use, frequency (n, Column%)								
≤ 1/day	15	10.6%	2	9.1%	3	7.7%	10	12.5%
> 1/day, ≤ 2.5/day	31	22.0%	4	18.2%	8	20.5%	19	23.8%
> 2.5/day, ≤ 4/day	40	28.4%	10	45.5%	8	24.2%	22	27.5%
> 4/day	45	31.9%	6	27.3%	19	57.6%	20	25.0%
Not ascertained	10	7.1%	0	0.0%	1	2.6%	9	11.3%
Heroin Ever Use (n, Column %)								
Yes	131	92.9%	22	100.0%	39	100.0%	70	87.5%
No	8	5.7%	0	0.0%	0	0.0%	8	10.0%
Not ascertained	2	1.4%	0	0.0%	0	0.0%	2	2.5%
Heroin Use, frequency (n, column %)								
Not regular user or not used in last 5 yrs	8	5.7%	0	0.0%	0	0.0%	8	10.0%
Not current user (used in last 5 yrs but not last yr)	5	3.5%	1	4.5%	1	3.0%	3	3.8%
< 1/week	6	4.3%	0	0.0%	2	5.1%	4	5.0%
≥ 1/week	120	85.1%	21	95.5%	36	92.3%	63	78.8%
Not ascertained	2	1.4%	0	0.0%	0	0.0%	2	2.5%
Cocaine use in last 5 years (n, Column %)								
Yes	117	83.0%	20	90.9%	35	89.7%	62	77.5%

	Total lung cancer cases with stage information		Local		Regional		Distant	
	(n = 141)		(n = 22, 15.6%)		(n = 39, 27.7%)		(n = 80, 56.7%)	
No	22	15.6%	2	9.1%	4	10.3%	16	20.0%
Not ascertained	2	1.4%	0	0.0%	0	0.0%	2	2.5%
Cocaine use in last 5 years, frequency (incl. speedball) (n, Column %)								
Non-user	22	15.6%	2	9.1%	4	1.2%	16	20.0%
< 1/week	37	26.2%	4	18.2%	12	3.5%	21	26.3%
≥ 1/week, < 1/day	36	25.5%	3	13.6%	12	3.5%	21	26.3%
≥ 1/day, < 3/day	24	17.0%	9	40.9%	4	1.2%	11	13.8%
≥ 3/day	20	14.2%	4	18.2%	7	2.1%	9	11.3%
Not ascertained	2	1.4%	0	0.0%	0	0.0%	2	2.5%

[a] Participants enrolled in NJ were asked about both their past smoking habits and their current smoking habits. Thirty-nine participants enrolled in NJ reported that they used to smoke in the past but no longer smoke. Participants enrolled in states other than NJ were only asked about their current smoking habits. Because the past smoker data was incompletely ascertained for the overall cohort, all NJ enrolled past smokers are grouped with participants who were not current smokers at the time of study enrollment for analysis.

[b] Alcohol use ever data was categorized based on if the subject had ever consumed alcohol in their lifetime. Alcohol consumption frequency data was categorized based on the following criteria: participants were classified as abstainers if they consumed ≤ 0.2 drinks per week; light drinkers if they consumed < 4 drinks per week; moderate drinkers if they consumed ≥ 4 drinks per week but ≤ 7 drinks per week for females or ≤ 14 drinks per week for males; and heavy drinkers if they consumed > 7 drinks per week for females or > 14 drinks per week for males.

Table 4
Standardized incidence ratios (SIRs) and standardized mortality ratios (SMRs) for all observed cancer cases (n = 245) and LC–caused deaths (n = 186, comparing observed cases in our cohort with expected cases. Expected values were calculated using individual risk estimates based on incidence rates from the New Jersey State Cancer Registry (NJSCR) and mortality rates from the Surveillance, Epidemiology, and End Results (SEER) program.

	Total Subject	OBS[a]	EXP[b]	SIR	95% CI		OBS[a]	EXP[b]	SMR	95%CI	
					Lower	Upper				Lower	Upper
Overall	9895	245	120.68	2.03	1.78	2.30	186	85.05	2.19	1.88	2.52
Sex											
Male	6371	172	89.18	1.93	1.65	2.24	132	65.67	2.01	1.68	2.38
Female	3524	73	31.50	2.32	1.82	2.91	54	19.39	2.79	2.09	3.63
Race											
NH White	3227	88	39.84	2.21	1.77	2.72	65	22.58	2.88	2.22	3.67
NH Black	3943	124	56.93	2.18	1.81	2.60	100	37.80	2.65	2.15	3.22
Hispanic	2665	32	23.62	1.35	0.93	1.91	20	24.23	0.83	0.50	1.27
Other Race	60	1	0.29	3.40	0.04	18.90	1	0.44	2.27	0.03	12.62
HIV status											
HIV Positive	2079	47	14	3.31	2.43	4.40	30	10.00	3.10	2.09	4.42
Prevalent HIV	1716	38	11.63	3.27	2.31	4.48	23	7.96	2.89	1.83	4.34
Incident HIV	363	9	2.56	3.51	1.60	6.66	7	1.73	4.05	1.62	8.34
HIV Negative	7759	196	105.40	1.86	1.61	2.14	156	74.68	2.09	1.77	2.44
Missing	57	2	1.08	1.85	0.21	6.68	0	0.69	0.00	0.00	5.32
Education Level											
< High School	4119	113	42.22	2.68	2.21	3.22	86	31.42	2.74	2.19	3.38
= High School	3594	74	45.08	1.64	1.29	2.06	56	31.00	1.81	1.36	2.35
> High School	2023	56	31.03	1.80	1.36	2.34	43	20.89	2.06	1.49	2.77
Missing	159	2	2.36	0.85	0.10	3.06	1	1.58	0.63	0.01	3.53
Smoking Frequency											
Non-current smoker (former and never)	820	5	11	0.44	0.14	1.02	4	7.91	0.51	0.14	1.29
< 1 pack/day	2566	49	30.88	1.59	1.17	2.10	38	22.68	1.68	1.19	2.30
1 - <1.5 pack/day	3883	97	47.79	2.03	1.65	2.48	78	33.62	2.32	1.83	2.90
≥ 1.5 pack/day	2409	93	27.80	3.35	2.70	4.10	65	19.00	3.42	2.64	4.36
Missing	217	1	2.78	0.36	0.00	2.00	1	1.84	0.54	0.01	3.02
Alcohol Use Frequency											
Abstain	3364	94	44.64	2.11	1.70	2.58	72	31.05	2.32	1.81	2.92
Alcohol user	6319	148	73.34	2.02	1.71	2.37	113	52.00	2.16	1.78	2.60
Light	2417	65	29.19	2.23	1.72	2.84	46	20.43	2.25	1.65	3.00
Moderate	1291	31	17.13	1.81	1.23	2.57	27	12.42	2.17	1.43	3.16
Heavy	2611	52	27.01	1.92	1.44	2.52	40	19.36	2.07	1.48	2.81

	Total Subject	OBS[a]	EXP[b]	SIR	95% CI		OBS[a]	EXP[b]	SMR	95%CI	
					Lower	Upper				Lower	Upper
Missing	212	3	2.70	1.11	0.22	3.24	1	1.79	0.56	0.01	3.11
Age at enrollment											
18–39 years old	7272	141	66.44	2.12	1.79	2.50	105	43.44	2.42	1.98	2.93
40–49 years old	2005	71	38.97	1.82	1.42	2.30	54	28.80	1.88	1.41	2.45
50 + years old	617	33	15.27	2.16	1.49	3.03	27	12.81	2.11	1.39	3.07
Missing	1	0	0.00	0.00	0.00	0.00	0	0.00	0.00	0.00	0.00
Smoking pack-years											
< 20	808	18	8.40	2.14	1.27	3.39	14	5.34	2.62	1.43	4.40
20–29	249	6	2.83	2.12	0.77	4.61	4	1.79	2.23	0.60	5.71
30+	267	9	3.20	2.82	1.29	5.35	5	2.08	2.4	0.77	5.61
Missing	8571	212	106.26	2.00	1.74	2.28	163	75.84	2.15	1.83	2.51

[a] OBS: observed number of LC cases.

[b] EXP: calculated expected number of LC cases.

Figures

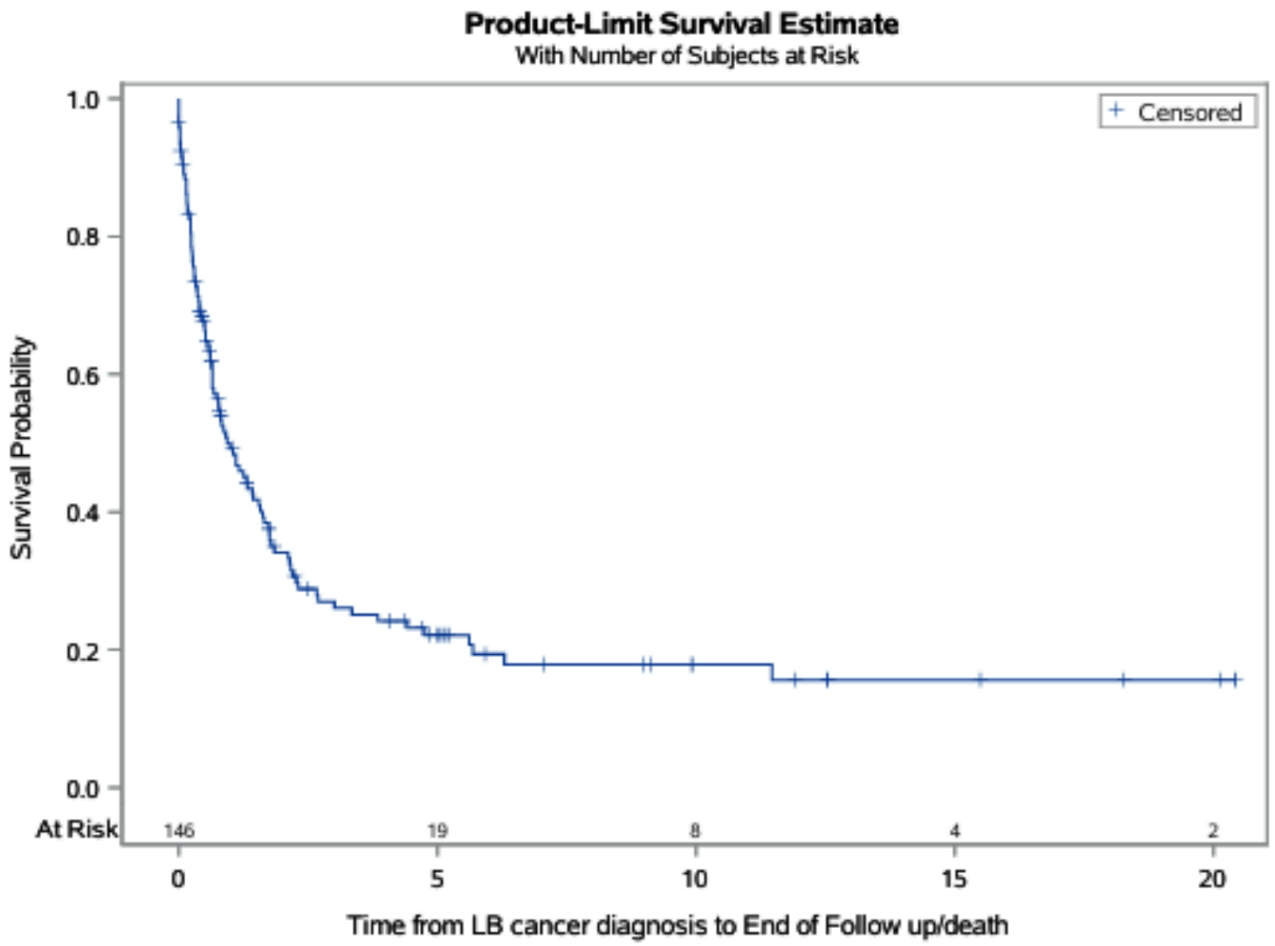

Figure 1

Kaplan–Meier survival curve from date of LC diagnosis to death due to LC (underlying or contributing cause) else censored alive. End-of-follow-up survival probability: 0.16.

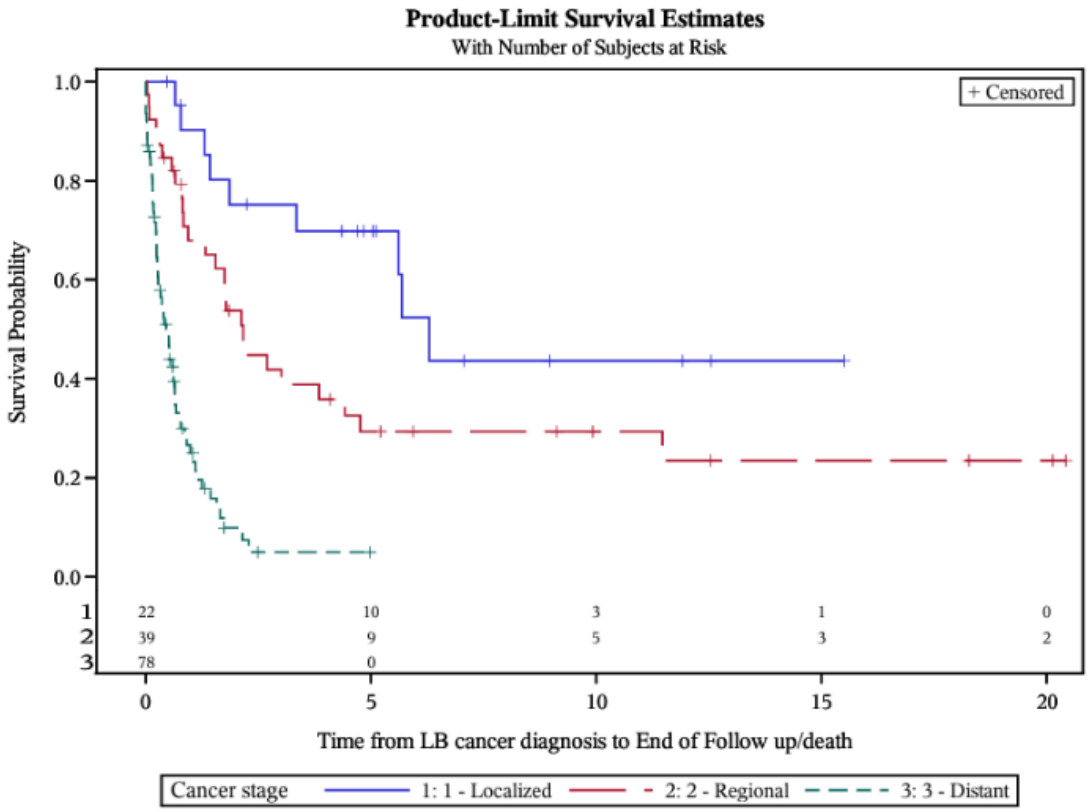

Figure 2

Kaplan–Meier survival curves for LC cases restricted to cases with stage at diagnosis information (including our reclassifications) and participants with LC as either the underlying or contributing cause of death. Survival time was calculated from the time of diagnosis of LC to the date of death due to LC or loss to follow-up. (Log-rank $p < 0.0001$; Wilcoxon $p < 0.0001$). End-of-follow-up survival probability: Localized, 0.44; Regional, 0.23; Distant, 0.05.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementalFigures.docx