



# Incidence, prevalence, and survival of lung cancer in the United Kingdom from 2000–2021: a population-based cohort study

George Corby<sup>1</sup>, Nicola L. Barclay<sup>1</sup>, Eng Hooi Tan<sup>1</sup>, Edward Burn<sup>1</sup>, Antonella Delmestri<sup>1</sup>, Talita Duarte-Salles<sup>2,3</sup>, Asieh Golozar<sup>4,5</sup>, Wai Yi Man<sup>1</sup>, Ilona Tietzova<sup>6</sup>, Daniel Prieto-Alhambra<sup>1,3^</sup>, Danielle Newby<sup>1</sup>

<sup>1</sup>Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Botnar Institute for Musculoskeletal Sciences, University of Oxford, Oxford, UK; <sup>2</sup>Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Barcelona, Spain; <sup>3</sup>Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands; <sup>4</sup>Odysseus Data Services, Cambridge, MA, USA; <sup>5</sup>OHDSI Center at the Roux Institute, Northeastern University, Boston, MA, USA; <sup>6</sup>First Department of Tuberculosis and Respiratory Diseases, First Faculty of Medicine, Charles University, Prague, Czech Republic

**Contributions:** (I) Conception and design: All authors; (II) Administrative support: A Delmestri, WY Man; (III) Provision of study materials or patients: D Prieto-Alhambra, A Golozar, I Tietzova; (IV) Collection and assembly of data: A Delmestri, WY Man; (V) Data analysis and interpretation: G Corby, E Burn, D Prieto-Alhambra, D Newby; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

**Correspondence to:** Daniel Prieto-Alhambra, PhD. Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Botnar Institute for Musculoskeletal Sciences, University of Oxford, Windmill Road, Oxford OX3 7LD, UK; Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands.

**Background:** Lung cancer is the leading cause of cancer-associated mortality worldwide. In the United Kingdom (UK), there has been a major reduction in smoking, the leading risk factor for lung cancer. Therefore, an up-to-date assessment of the trends of lung cancer is required in the UK. This study aims to describe lung cancer burden and trends in terms of incidence, prevalence, and survival from 2000–2021, using two UK primary care databases.

**Methods:** We performed a population-based cohort study using the UK primary care Clinical Practice Research Datalink (CPRD) GOLD database, compared with CPRD Aurum. Participants aged 18+ years, with 1-year of prior data availability, were included. We estimated lung cancer incidence rates (IRs), period prevalence (PP), and survival at 1, 5 and 10 years after diagnosis using the Kaplan-Meier (KM) method.

**Results:** Overall, 11,388,117 participants, with 45,563 lung cancer cases were studied. The IR of lung cancer was 52.0 [95% confidence interval (CI): 51.5 to 52.5] per 100,000 person-years, with incidence increasing from 2000 to 2021. Females aged over 50 years of age showed increases in incidence over the study period, ranging from increases of 8 to 123 per 100,000 person-years, with the greatest increase in females aged 80–89 years. Alternatively, for males, only cohorts aged over 80 years showed increases in incidence over the study period. The highest IR was observed in people aged 80–89 years. PP in 2021 was 0.18%, with the largest rise seen in participants aged over 60 years. Median survival post-diagnosis increased from 6.6 months in those diagnosed between 2000–2004 to 10.0 months between 2015–2019. Both short and long-term survival was higher in younger cohorts, with 82.7% 1-year survival in those aged 18–29 years, versus 24.2% in the age 90+ years cohort. Throughout the study period, survival was longer in females, with a larger increase in survival over time than in males.

**Conclusions:** The incidence and prevalence of lung cancer diagnoses in the UK have increased, especially in female and older populations, with a small increase in median survival. This study will enable future comparisons of overall disease burden, so the overall impact may be seen.

**Keywords:** Lung cancer; incidence; prevalence; cancer survival

<sup>^</sup> ORCID: 0000-0002-3950-6346.

Submitted Mar 18, 2024. Accepted for publication Jul 03, 2024. Published online Sep 21, 2024.

doi: 10.21037/tlcr-24-241

View this article at: <https://dx.doi.org/10.21037/tlcr-24-241>

## Introduction

Lung cancer is the leading cause of cancer-associated mortality worldwide, with 1.8 million deaths in 2020 (1). New diagnoses are predicted to nearly double by 2070 meaning it will continue to be a major cause of morbidity and mortality globally (2).

As early as 1962, the British Royal College of Physicians (RCP) published *Smoking and Health* (3), which established a clear and important link between smoking and lung cancer. Although smoking remains the greatest risk factor for all lung cancer subtypes, other factors such as family history (4), and exposure to arsenic (5), radon (6), biomass fuels, asbestos (7), or a broad range of occupational chemicals, have been identified as increasing risk. These risk factors can influence the proportion of patients with each subtype of lung cancer (2).

The United Kingdom (UK) Office of National Statistics (ONS) surveys indicate that 60 years from the initial RCP

report, smoking prevalence in the UK continues to decrease, from 45% in the early 1970s to 14% in 2020 (8), with a concurrent rise in the use of electronic cigarettes (9,10). This reduction in the prevalence of smoking has coincided with a revolution in treatments of lung cancer, with major advancements in surgery (11), targeted drug treatments (12), and a shift towards multidisciplinary team (MDT) management (13). Together, this has corresponded with a major fall in mortality rates of lung cancer, with a 38% fall in mortality rate per 100,000 from the mid-1980s to late-2010s (13). This reduction is driven by a decrease in male mortality. Female mortality has slightly increased from 1970, peaking in 2010, despite the prevalence of smoking falling in both sexes (13).

In 2023, the UK introduced targeted screening for lung cancer. People aged 55–74 years with a general practice (GP) health record documenting a smoking history will be invited to interview for a risk assessment, after which they may be offered a low-dose computed tomography (CT) scan. When the rollout is complete, it is estimated 325,000 people will become newly eligible for a scan every year. Focussing on those with the highest risk will enable earlier diagnosis and potentially better survival, with reduced iatrogenic harm from screening-related radiation exposure (14,15).

Due to changes in risk factor exposure, and the introduction of the new screening in 2023, a comprehensive assessment of the trends of lung cancer in different population strata using routinely collected data from primary care, is required in the UK. Understanding these trends in lung cancer is an important aspect of population healthcare planning, particularly considering the introduction of screening high-risk individuals. Therefore, the aim of this study is to describe lung cancer burden and trends in terms of incidence, prevalence, and survival from 2000–2021 using two large and representative primary care databases from the UK. We present this article in accordance with the STROBE reporting checklist (available at <https://tlcr.amegroups.com/article/view/10.21037/tlcr-24-241/rc>).

## Methods

### *Study design, setting, and data sources*

We carried out a population descriptive cohort study using routinely collected primary care data from the UK. People

### Highlight box

#### Key findings

- The incidence and prevalence of lung cancer diagnoses in the United Kingdom (UK) have increased, especially in female and older populations, with a small increase in median survival. Both short- and long-term lung cancer survival has improved over time for both sexes.

#### What is known and what is new?

- Lung cancer is the leading cause of cancer-associated mortality worldwide. In the UK, there has been a major reduction in smoking, the leading risk factor for lung cancer, alongside advances in lung cancer treatments, and changes in baseline population demographics.
- With the introduction of the UK lung cancer screening programme in 2023, an up-to-date assessment of the trends of lung cancer before the introduction of this screening programme is required.

#### What is the implication, and what should change now?

- This study will help to enable future comparisons of overall disease burden, and changes in trends in the incidence, prevalence, and survival with lung cancer, so the impact of screening alongside novel treatments, and changing demographics may be seen.
- Further research is required to understand increasing disease burden in females.

with a diagnosis of lung cancer and a background cohort (denominator population) were identified from Clinical Practice Research Datalink (CPRD) GOLD to estimate overall survival, incidence, and prevalence. We additionally carried out this study using CPRD Aurum to compare the results for GOLD. Both these databases contain pseudonymised patient-level information on demographics, lifestyle data, clinical diagnoses, prescriptions, and preventive care provided to patients and collected by the National Health Service (NHS) as part of their care and support. CPRD GOLD contains data from across the UK, from GP practices using the Vision<sup>®</sup> software system, whereas Aurum only contains data from England, from GP practices using the EMIS<sup>®</sup> software system. Both databases are established primary care databases covering over 50 million people (16), and both were mapped to the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) (17,18). Uniquely, the mapping of both to a common data model (OMOP) allowed us to analyse both simultaneously, and using the same exact analytical code. The use of CPRD data in this study was approved by the Independent Scientific Advisory Committee (No. 22\_001843). The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Individual consent for this retrospective analysis was waived.

### *Study participants and time at risk*

All participants were required to be aged 18 years or older and have at least 1-year of prior history. For the incidence, prevalence, and survival analyses, the study cohort consisted of individuals present in the database from 1<sup>st</sup> January 2000. For CPRD GOLD, these individuals were followed up to whichever came first: practice stopped contributing to the database, patient left the practice, date of death, or the 31<sup>st</sup> of December 2021 (the end of the study period) whereas for Aurum, the end of the study period was 31<sup>st</sup> of December 2019. For the survival analysis, only individuals with newly diagnosed lung cancer were included. Any patients whose death and cancer diagnosis occurred on the same date were removed from the survival analysis.

### *Lung cancer definition*

We used Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) diagnostic codes to identify lung cancer events. Diagnostic codes indicative of either non-malignant cancer or secondary metastases from other

organs, were excluded as well as diagnosis codes indicative of melanoma and lymphoma occurring in the organs of interest. The study outcome cancer definition was reviewed with the aid of the CohortDiagnostics R package (19). This package was used to identify additional codes of interest and to remove those highlighted as irrelevant based on feedback from clinicians with oncology, primary care, and real-world data expertise through an iterative process during the initial stages of analyses. The clinical code lists used to define lung cancer can be found in [Table S1](#).

OMOP-based computable phenotypes are available, together with all analytical code at our GitHub repository to enable reproducibility (see statistical methods). For survival analyses, mortality was defined as all-cause mortality based on CPRD GOLD date of death records, which have been previously validated and shown to be over 98% accurate (20).

### *Statistical analysis*

The population characteristics of patients with a diagnosis of lung cancer were summarised, with median and interquartile range (IQR) used for continuous variables and counts and percentages used for categorical variables.

We calculated the overall and annualised crude incidence rates (IRs) and annualised prevalence for lung cancer from 2000 to 2021. For incidence, the number of events, the observed time at risk, and the IR per 100,000 person-years were summarised along with 95% confidence intervals (95% CI). Annual IR were calculated as the number of incident lung cancer cases as the numerator and the number of person-years in the general population within that year as the denominator whereas overall incidence was calculated from 2000 to 2021.

Age-standardized IRs were calculated using the 2013 European Standard Population (ESP2013) (21). The ESP2013 serves as a standard population with a predefined age distribution where to account for differences in age structures between different populations to ensure fair comparisons. The ESP2013 provides predefined age distribution in 5-year age bands; therefore, we collapsed these to obtain distributions for 10-year age bands used in this study. We used the age distribution of 20–29 years from ESP2013 for age-standardization as age distributions were not available for 18–29 years age band used in this study.

Period prevalence (PP) was calculated on 1<sup>st</sup> January for the years 2000 to 2021, with the number of patients fulfilling the case definition for lung cancer as the

numerator. The denominator was the number of patients on 1<sup>st</sup> January in the respective years for each database. The number of events, and prevalence (%) were summarised along with 95% CIs.

For survival analysis, we used the Kaplan-Meier (KM) method to estimate the overall survival probability from observed survival times with 95% CIs. We estimated the median survival and survival probability 1, 5, and 10 years after diagnosis.

All results were stratified by database, by age (10-year age bands apart from the first and last age bands which were 18 to 29 years and 90 years and older respectively) and by sex. Additionally, for GOLD only, we stratified by calendar time of cancer diagnosis (2000–2004, 2005–2009, 2010–2014, 2015–2019 and 2020–2021) to understand if survival has changed over time. To avoid re-identification, we do not report results with fewer than five cases.

For Aurum, the same statistical analyses were performed using data from 1<sup>st</sup> January 2000 to 31<sup>st</sup> December 2019 to compare the results obtained from GOLD.

The statistical software R version 4.2.3 was used for analyses. For calculating incidence and prevalence, we used the IncidencePrevalence R package (22). For survival analysis, we used the survival R package (23). All analytic code used to perform the study is available at <https://github.com/oxford-pharmacoepi/EHDENcancerIncidencePrevalence>.

## Results

### *Patient populations and characteristics*

Overall, there were 11,388,117 eligible patients identified from January 2000 to December 2021 from CPRD GOLD. Attrition tables can be found in the [Table S2](#). A summary of baseline patient characteristics of those with a diagnosis of lung cancer is shown in [Table 1](#). Further stratifications of patient characteristics by UK region, and smoking status for CPRD GOLD can be found in [Tables S3,S4](#).

Overall, from the 45,563 patients with a diagnosis of lung cancer, patients were more likely to be male (54%), with a median age of 72 years old. The highest percentage of patients were aged 70–79 years old, contributing to 36.7% of diagnosed patients. Patients with lung cancer had a high prevalence of chronic obstructive pulmonary disease (COPD) (25%) as well as cardiovascular comorbidities such as heart disease (23.5%) and hypertensive disorder (27.2%) at presentation. Similar observations were seen across both databases. A similar table with detailed baseline

characteristics for Aurum patients is available in [Table S5](#).

For GOLD, stratification by UK region showed characteristics were generally similar across the different regions apart from smoking status which had lower proportions for England compared to the other regions ([Table S3](#)). Stratification on smoking status showed larger proportions of younger patients (40–59 years) who were smokers compared to non and former smokers, with the average age at presentation being 8 years older in non-smokers than smokers, 77 versus 69 respectively ([Table S4](#)). Smokers had higher proportions of diagnoses of COPD, depressive disorders, and peripheral vascular disease compared with the non-smoking group.

### *IRs stratified by calendar year, age, and sex*

The overall IR of lung cancer from 2000 to 2021 was 52.0 per 100,000 person-years (95% CI: 51.5 to 52.5) for GOLD. Females had lower overall IR [47.2 per 100,000 person-years (95% CI: 46.5 to 47.8)] compared to males [57.0 per 100,000 person-years (95% CI: 56.3 to 57.7)], with similar rates in Aurum. Annualised IRs increased across both databases ([Figure 1](#)). Females showed increasing IRs over the study period, while males showed a more stable trend in both databases. Age standardized results using the ESP2013 for CPRD GOLD show, after 2004, a decreasing trend in incidence for males, whereas an increase in incidence over the study period for females ([Figure S1](#)). Further stratification by UK region showed a similar trend for males and females ([Figure S2](#)). All study results can be found in and downloaded from an interactive web application: <https://dpa-pde-oxford.shinyapps.io/LungCancerIncPrevSurvShiny/>.

Overall IRs increased with age up to 80–89 years across both databases ([Table S6](#)). Annualised IRs for each age group ([Figure 2](#)) over time showed for those aged 70–89 years a gradual increase between 2004–2019. Whereas for those 50–59 years of age there was a gradual decline in IRs since 2004. For those aged 60–69 years, IRs were stable from 2004–2019 for GOLD, whereas for Aurum, IRs increased from 2000 to 2013 before stabilising in 2019. For those over 90 years of age, there was also a stabilisation of IRs from 2014 with larger differences between the databases. Younger patients (30–49 years of age) showed relatively stable IRs over the study period.

Stratification by age and sex ([Figure S3](#)) showed similar trends across both databases. For females aged 60–89 years, IRs increased between 2000–2019, whereas for males, IRs

**Table 1** Baseline characteristics of lung cancer patients at the time of diagnosis for CPRD GOLD

Database	CPRD GOLD
Number of lung cancer patients	45,563
Sex: male, N (%)	24,569 (53.9)
Age (years), median [IQR]	72 [65 to 79]
Age groups (years), N (%)	
18–29	24 (0.1)
30–39	129 (0.3)
40–49	1,030 (2.3)
50–59	4,660 (10.2)
60–69	12,249 (26.9)
70–79	16,745 (36.8)
80–89	9,546 (21.0)
90+	1,180 (2.6)
Prior history (days), median [IQR]	3,660 [2,009 to 5,352]
Smoking status (any time 5 years prior), N (%)	
Non-smoker	7,154 (15.7)
Former smoker	543 (1.2)
Current smoker	22,019 (48.3)
Missing	15,847 (34.8)
General conditions (any time prior), N (%)	
Atrial fibrillation	3,207 (7.0)
Cerebrovascular disease	3,840 (8.4)
Chronic liver disease	247 (0.5)
Chronic obstructive lung disease	11,163 (24.5)
Coronary arteriosclerosis	677 (1.5)
Crohn’s disease	154 (0.3)
Dementia	772 (1.7)
Depressive disorder	6,397 (14.0)
Diabetes mellitus	5,321 (11.7)
Gastroesophageal reflux disease	1,261 (2.8)
Gastrointestinal haemorrhage	3,113 (6.8)
Heart disease	10,704 (23.5)
Heart failure	2,006 (4.4)
HIV	23 (0.1)
Hyperlipidemia	4,586 (10.1)
Hypertensive disorder	12,404 (27.2)
Ischemic heart disease	6,237 (13.7)

**Table 1** (continued)

**Table 1** (continued)

Database	CPRD GOLD
Obesity	928 (2.0)
Osteoarthritis	9,841 (21.6)
Peripheral vascular disease	2,260 (5.0)
Pneumonia	2,549 (5.6)
Pulmonary embolism	804 (1.8)
Renal impairment	6,156 (13.5)
Ulcerative colitis	181 (0.4)
Venous thrombosis	2,331 (5.1)

CPRD, Clinical Practice Research Datalink; IQR, interquartile range; HIV, human immunodeficiency virus.

were relatively stable across this period. Males had higher IRs compared with females apart from those 30–59 years of age, where IRs were similar between sexes. Also, similar IRs were seen for females and males aged 60–69 years from 2015 onwards for GOLD.

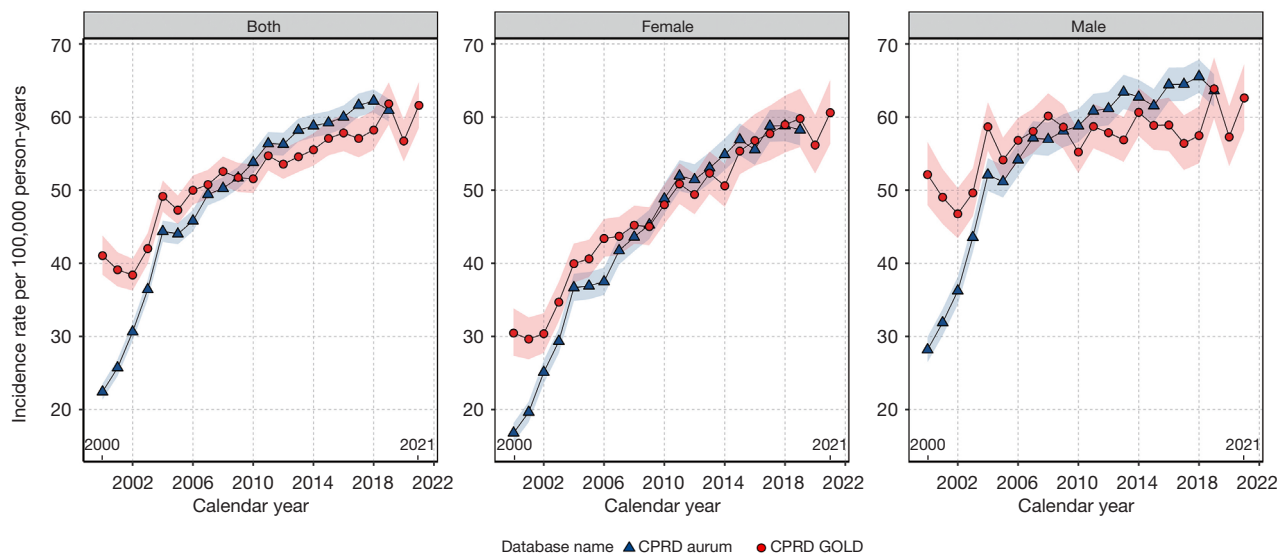
***PP for study population with database, age, and sex stratifications***

For the whole population, the PP for lung cancer in 2021 was 0.18% (95% CI: 0.17% to 0.18%) for GOLD. PP in 2019 was similar across both databases (~0.17%). Sex stratification showed PP in 2019 were slightly higher in females [0.175% (95% CI: 0.168% to 0.181%)] compared to males [0.156% (95% CI: 0.150% to 0.162%)] in GOLD with smaller differences in Aurum. In GOLD, PP has increased 2.7-fold from 2000 to 2021, with females having a larger fold increase (3.7-fold) in PP across the study period compared to males (2.0-fold), with similar observations seen in Aurum (Figure 3). Furthermore, stratification by UK region for GOLD showed similar trends across the different regions (Figure S4).

When stratifying by age group, PP in 2019 was highest in those 80–89 years of age (0.80% Aurum, 0.72% GOLD) with this age group seeing the largest change in PP over the study period (Figure 4). Overall, most age groups showed increases in PP over the study period for both databases.

Stratification by sex and age showed similar trends between males and females across age groups (Figure S5). For those aged 50–59 years of age, PP for males remained relatively stable over the study period from 2005, whereas PP for females increased. For those 60–69 years of age, PP





**Figure 1** Annualised incidence rates for lung cancer from 2000 to 2021 stratified by database and sex. Bands show 95% confidence interval. CPRD, Clinical Practice Research Datalink.

increased for females from 2010, whereas for males in this age group PP initially increased from 2010 to then decrease from 2014 in GOLD (in Aurum PP continued to increase, albeit at a slower rate than in females). Overall, males had higher PP compared to females for those aged 80 years and older. For those aged 50–79 years males had higher PP earlier in the study period with females having higher PP towards the end of the study period. For those aged 40–59 years, there were no differences in PP over time between males and females.

#### **Overall survival rates for the cancer population with age, sex, and calendar year stratification**

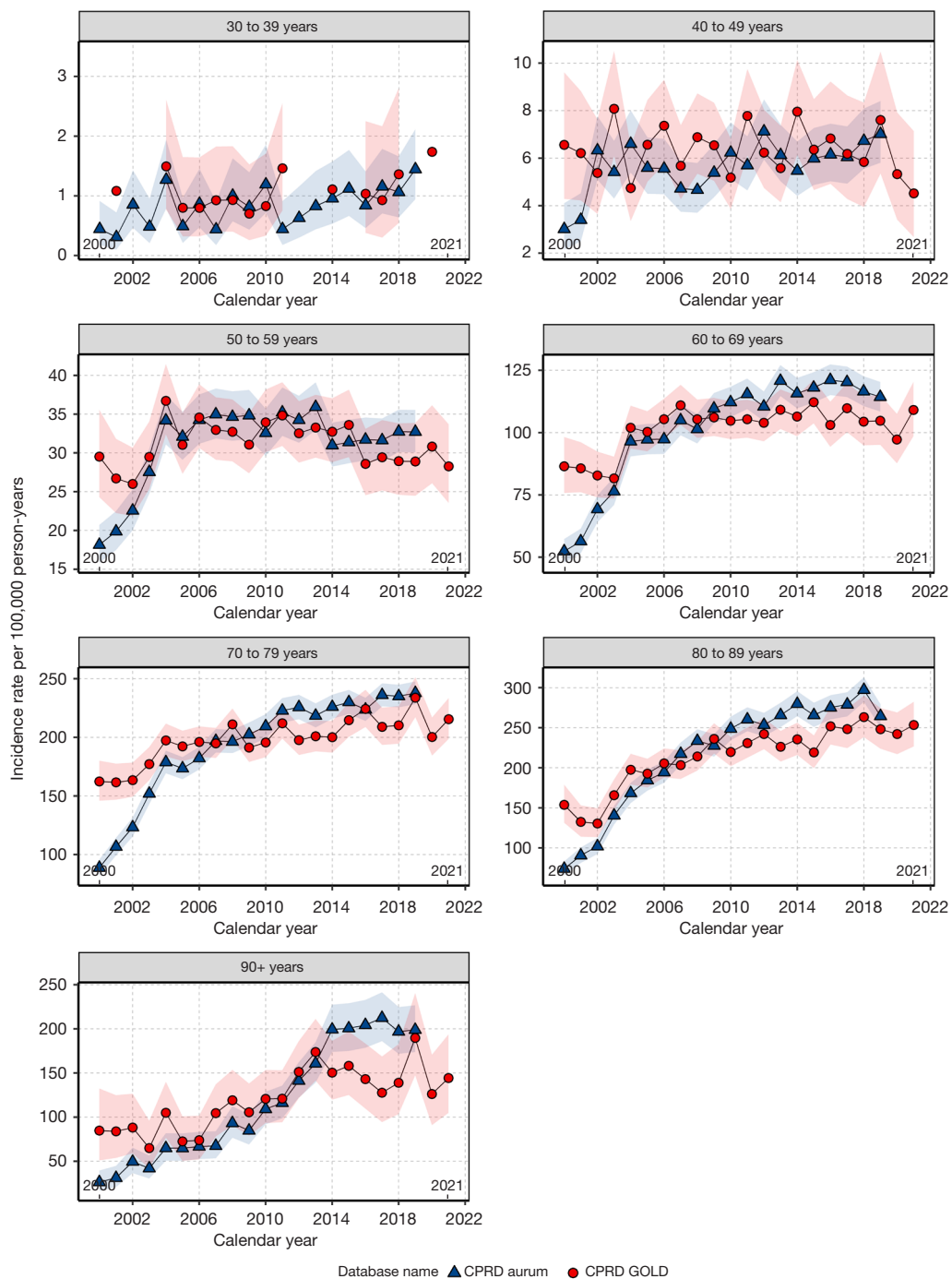
For GOLD there were 43,903 patients with 35,381 deaths (80.6% of patients) over the study period with a median follow-up of 0.54 years (IQR, 0.18–1.39 years). For Aurum, there were 86,710 patients with 67,421 deaths (77.8% of patients) over the study period with a median follow-up of 0.58 years (IQR, 0.19–1.49 years).

Figure S6 shows survival curves for the overall population and stratified by sex. The median survival for the whole population was 0.66 years (95% CI: 0.64–0.67) and 0.71 years (95% CI: 0.70–0.72) in GOLD and Aurum respectively. Survival after 1, 5 and 10 years after diagnosis was 39.0%, 12.0% and 6.4% for GOLD, with similar results across the different UK regions (Figure S7) and similar values for Aurum.

Median survival was higher in females (0.75–0.81 years) compared to males (0.60–0.64 years) across both databases. Regarding short- and long-term survival, females had better survival compared with males (Table S7). When stratifying by age group, median survival decreased with age for both databases from 1.3–1.5 years for those aged 30–39 years to ~0.35 years for those aged 90 years and older (Table S8). Survival at 1, 5, and 10 years also decreased with increasing age (Table S9).

Median survival increased from 6.6 months in those diagnosed between 2000 to 2004 to 10 months for those diagnosed between 2015 to 2019 for both sexes (Figure 5, Table S10). Median survival was similar for those diagnosed in 2020–2021 to 2015–2019. For different age groups, median survival increased over time for most age groups (40–89 years of age) apart from those 90 years and older where median survival has not changed over time.

Overall, 1-year survival has increased between 2000–2004 and 2015–2019 for the whole population with similar trends for both sexes for GOLD. For the whole population, 1-year survival increased from 33% in 2000–2004 to 45% in 2015–2019. For 5-year survival, there were linear increases in survival over time for GOLD from 2000–2004 to 2015–2019 (Table S11). When stratifying by age group, both 1- and 5-year survivals were higher in those diagnosed between 2015–2019 compared to those diagnosed between 2000–2004 for those 50–89 years of age with similar patterns for both sexes.

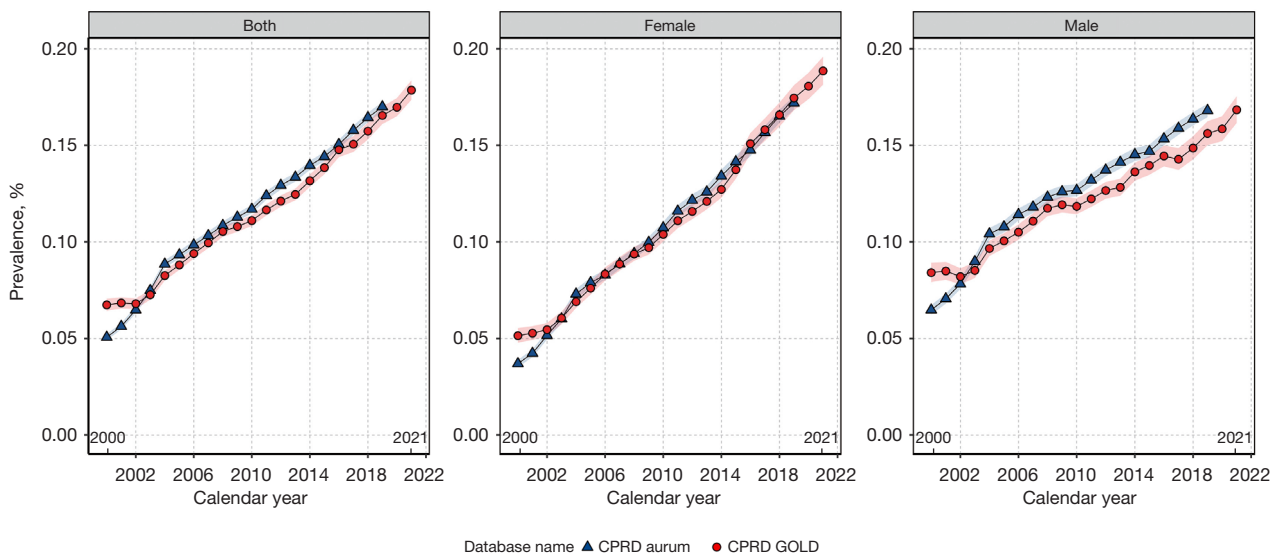


**Figure 2** Annualised incidence rates from 2000 to 2021 stratified by database and age group. Bands show 95% confidence interval. CPRD, Clinical Practice Research Datalink.

**Discussion**

This study provides a comprehensive analysis of trends in lung cancer incidence, prevalence, and survival in the UK.

In this large cohort of over 11 million people, the incidence and prevalence of lung cancers in the UK has increased from 2000 to 2021, while both short- and long-term survival has slightly improved across all age groups over this



**Figure 3** Annual period prevalence from 2000 to 2021 for the whole population and stratified by sex. Bands show 95% confidence interval. CPRD, Clinical Practice Research Datalink.

time period, with better survival in females than males.

IRs reported here are broadly in line with those estimated from analyses from the Rapid Cancer Registration Databases however are lower than National Cancer Statistics (24-26). However, it must be noted that National Cancer Statistics include cancers of the lung/bronchus and trachea whereas this study and the Rapid Cancer Registration Databases only include cancers of the lung/bronchus. For age and sex, studies have reported higher IRs in males compared to females and higher incidence with increasing age peaking in those aged 80 to 89 years of age in line with our estimates (26,27). Other studies have also shown increases in prevalence, particularly in females, where prevalence has overtaken males in recent years (1,28).

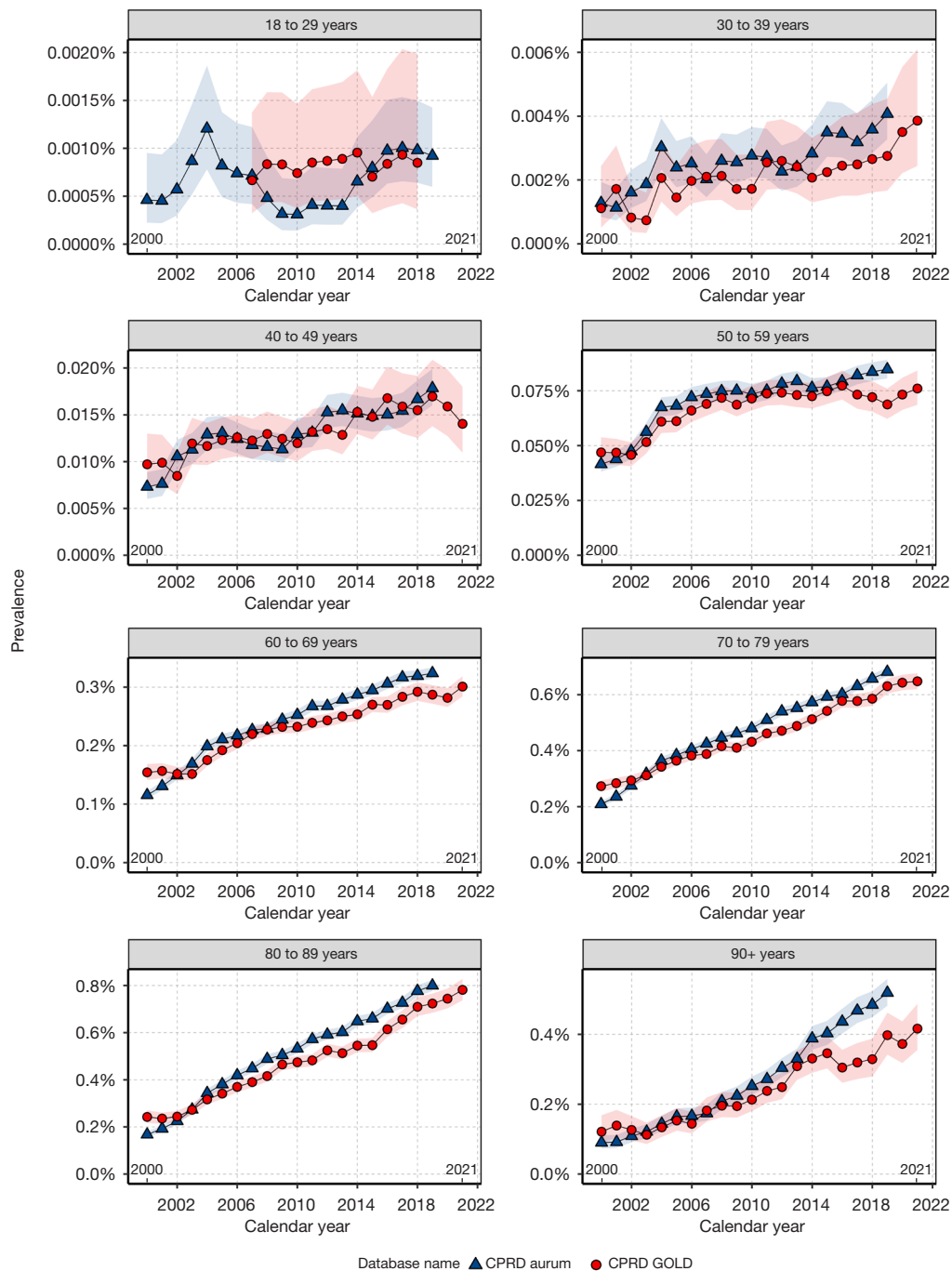
The increase in incidence and prevalence of lung cancer over time has been reported particularly in countries with higher levels of economic development as well as countries with higher smoking prevalence and air pollution (1,2,29). These increases have largely been driven by females. However, other studies show a decline or stabilisation of the incidence of lung cancer due to decreases in males (1,30,31). Our age standardized IRs to the European Standard Population 2013 are in line with these observations for females and males. For this work we report crude IRs trends overall and stratified by sex and age to understand health provision in the UK. IRs are known to vary by age as shown by this work and this is concordant with national cancer

statistics (25) which show decreases in incidence over time in those aged 50–59 and aged over 80 years particularly after 2014 with increases in those aged 60–79 years. These age differences in time trends are likely due to birth cohort effects as well as differences in modifiable risk factor exposures (25). However, other studies also show IRs have decreased over time in all age groups which could be driven by a larger decrease in male incidence over time even with the increasing incidence in females (2,32).

Encouragingly, despite the increases in incidence and prevalence, our study shows 1- and 5-year survival has improved over the past 20 years, with increases of 11% and 7% respectively. Our results are in line with the 2023 National Lung Cancer Audit for England and Wales (33), UK cancer registry (34), as well as other international studies (1,35-37). However, despite these improvements, short- and long-term survival is still low compared with other common cancers such as breast and prostate. Interestingly, 5-year survival rates in this study are broadly comparable across Europe with values between 10–20% (1,35-37).

The rise in lung cancer could be for numerous reasons, making future trends difficult to predict. Smoking is the greatest risk factor for the disease, with a 20-year delay between smoke exposure and cancer onset (38). In our study, a high proportion of lung cancer patients (74%) were either current or ex-smokers; which is approximately five times higher than the overall population smoking prevalence (8) but slightly lower than UK national audit data estimates of

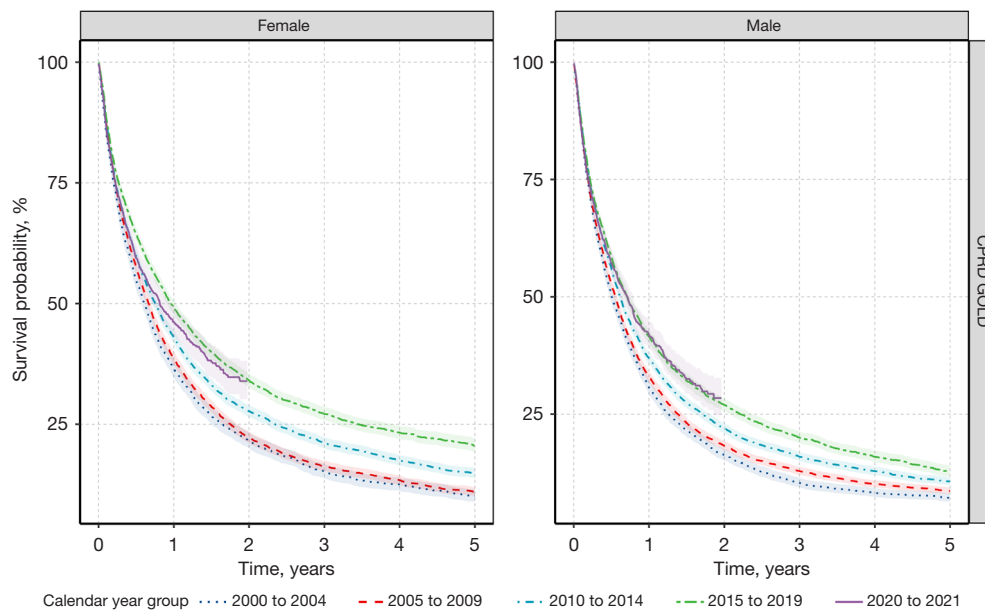




**Figure 4** Annual prevalence from 2000 to 2021 stratified by database and age group. Bands show 95% confidence interval. CPRD, Clinical Practice Research Datalink.

smoking status in people with lung cancer (33). However, following numerous successful public health interventions, the prevalence of smoking has fallen in the UK (8) as well as many other countries across Europe (39), although may

have increased during recent lockdowns (40). Furthermore, UK legislation has reduced the impact of 'second-hand' smoking which has amplified the impact of the reduction in smoking prevalence further in the UK (41). Therefore,



**Figure 5** Kaplan-Meier survival curve of lung cancer stratified by sex and calendar year of diagnosis (2000–2004, 2005–2009, 2010–2014, 2015–2019 and 2020–2021). Bands show 95% confidence interval. CPRD, Clinical Practice Research Datalink.

increases in disease burden in this study cannot be attributed to smoking prevalence alone. However, in our study, we further see an overall shift, from 2001–2021, in the age distribution of incident diagnoses, from younger to older participants—with incidence rising especially in females aged over 50 years, and males over 80 years. This may reflect that lower exposure to cigarette smoke is causing people that are predisposed to developing lung cancer, to develop it at an older age.

Alternatively, in recent years, electronic cigarettes have been promoted as a harm reduction strategy to a greater extent than in many other countries. The effects of this are largely unknown, especially with the lag time before cancer onset, so the effect of this may not yet be full seen (42).

Other risk factors include COPD, which is an independent risk factor for lung cancer (2). Nearly a quarter (24.5%) of patients had a prior diagnosis of COPD across both sexes, which is higher than existing estimates of COPD prevalence in the overall population (43).

In never-smokers, defined as individuals who have smoked fewer than 100 cigarettes in their lifetime (39), radon is the leading environmental cause of lung cancer (44). As with smoking, there is a substantial exposure–cancer–mortality delay (45), and public health campaigns have aimed to reduce exposure (6). Air pollution is also an established risk factor for lung cancer (46) with a study using UK Biobank

demonstrating additive interactions between air pollution and genetic-related risk factors for lung cancer (47). Despite apparent reductions of exposure to the discussed major modifiable risk factors in the UK, incidence and prevalence of lung cancer has increased in our study.

While not a sex-specific cancer, lung cancer shows sex-specific trends in our data as well as has been reported nationally and internationally (48–50). A recent review attributed this finding to many factors such as the slower relative reduction in female smoking compared to males (8,49), coupled with females being exposed to different risk factors, as well as differences in oestrogen levels (49), exposure to human papillomavirus (HPV) (51) and genetic polymorphisms (49). Furthermore, females have been reported to have a higher frequency of mutations in critical genes, such as TP53 and the KRAS leading to higher risk of the disease (52–54).

Although incidence and prevalence are increasing in females, overall survival is better in females compared to males in line with a similar study using primary care data from the UK (27). Again, reasons for this difference are likely to be multifactorial involving different risk factors, treatment decisions, and cancer histology (48). Furthermore, changes in age and sex distributions of those diagnosed with lung cancer over time could also contribute to the observed differences in survival. Survival from lung cancer has

substantially improved since 2000, which has corresponded with major advancements in lung cancer treatments (2,12). Towards the start of this period, management of this disease progressively included a multi-disciplinary team (MDT) for cancer care (11,55,56). Additionally, lung cancer treatments have further shifted towards more targeted regimens, after demonstrating better survival than existing therapy in trials (12). Additionally, there has been a concurrent expansion of lung cancer surgery, which has been shown to further improve treatment outcomes (57,58).

The increases in survival, incidence, and prevalence, in this study could also be due to better diagnostic methods and improved public awareness of lung cancer symptoms, leading to earlier detection. For instance, a 2012 UK Department of Health campaign was implemented to raise awareness of persistent cough as a lung cancer symptom, leading to a 3.1% increase in the proportion of non-small cell lung cancer (NSCLC) diagnosed at stage 1 (59), with similar campaigns since (60). If diagnoses are indeed occurring at an earlier stage, lead-time bias may result in improvements in the survival data that do not exist in practice due to the fact the cancer was simply detected earlier, even if the treatment given and ultimate date of death is unchanged (61). Of course, it may be the case that whilst cancers are diagnosed sooner, this is coupled with better treatment, not only for an equivalent stage due to improved therapy but also because even better treatment may be delivered at the earlier stage of cancer. The introduction of CT screening may increase the recorded incidence, prevalence, and survival, of lung cancer for similar reasons, which should be explored by future research.

The main strength of this study is the use of a large primary care database covering the whole of the UK and validation of the results using another database from England. CPRD GOLD covers primary care practices from England, Wales, Scotland, and Northern Ireland whereas CPRD Aurum covers primary care practices in England. The similarity between the results in both databases, and their overall agreement with National Cancer Statistics (NCS) and national audit programmes provide increased generalizability across the UK. The sharp increase in incidence at the start of the study period between 2000–2004 could be due and the institution of cancer quality improvement measures by the NHS in 2003 as well as the introduction of the Quality and Outcomes Framework (QOF) in 2004 which encourages general practitioners to record all new cases of cancer which could partly explain

the increase in cancer recording (62).

Another strength of our study is the inclusion of a complete study population database for the assessment of incidence and prevalence. In contrast, cancer registry studies extrapolate the registry data to the whole population using national population statistics, potentially introducing biases (63,64). The high validity and completeness of mortality data with over 98% accuracy compared to national mortality records (20) allowed us to examine the impact of calendar time on overall survival—one of the key outcomes in cancer care.

Our study also has some limitations. Firstly, we used primary care data without linkage to a cancer registry which could lead to misclassification, delayed recording, or incompleteness of cancer diagnoses (65) which could result in lower estimates. A previous validation study has shown high accuracy and completeness of cancer diagnoses in primary care records. However, although a high positive predictive value (>80%) was observed, a lower sensitivity was also reported (66). Secondly, our use of primary care records, means we do not have information on tumour histology, genetic mutations, staging, specific treatments, and further environmental factors, which means that survival estimates might not fully account for variability in outcomes based on these factors (67). Therefore, our survival estimates may overestimate survival in those with higher staging as well as those with specific subtypes or mutations associated with poorer survival such as small cell lung cancer (SCLC) (68). Other factors, such as socio-economic status, environmental exposures such as air pollution and ethnicity could also result in different values for incidence, prevalence, and survival (46,47). Thirdly, in this study, we calculated overall survival, which does not differentiate between deaths caused by cancer *vs.* other causes. Therefore, it is a broad measure of overall survival rather than specifically cancer mortality. However, with the introduction of low-dose CT screening being introduced in the UK for targeted groups, the use of overall survival will enable the assessment of how much screening promotes overall survival and prevents all-cause mortality (including for instance changes related to non-lung cancer incidental diagnoses on CT screening), not just mortality related to lung cancer, and the overall benefit (or harm) associated with lung cancer treatments. Finally, smoking status was missing in 34.8% of lung cancer patients in this study with 48.3% of patients recorded as smokers which is around 3–5× the overall population prevalence of smoking for this period, however may be higher (8). Although not possible

in CPRD due to missingness which could create biases in results, future research using alternative databases could be used to quantify the relative contribution of smoking in 'pack-years', alongside alternative discussed risk factors.

## Conclusions

Despite the falling prevalence of smoking in the UK, the incidence and prevalence of lung cancer is increasing. Reassuringly, the improvements in survival over the study period highlight the development of better-targeted treatments and earlier diagnosis of high-risk populations. However, the rise of lung cancer even with a decrease in smoking is a cause for concern, particularly in females. Further work needs to focus on understanding this demographic shift, to explain why lung cancer continues to rise, which could lead to better prevention, earlier diagnosis, and further targeted treatments. As the UK Lung Cancer Screening Programme is introduced, with a target of 100% implementation by 2030, our study has the potential to enable subsequent comparisons of not only survival rates, but the baseline medical characteristics of people at diagnosis ultimately enabling a more comprehensive assessment of the impact of the screening programme and resulting public health interventions in the UK.

## Acknowledgments

*Funding:* This activity under the European Health Data & Evidence Network (EHDEN) and OPTIMA has received funding from the Innovative Medicines Initiative 2 (IMI2) Joint Undertaking under grant agreement No. 806968 and No. 101034347 respectively. IMI2 receives support from the European Union's Horizon 2020 research and innovation programme and European Federation of Pharmaceutical Industries and Associations (EFPIA). The sponsors of the study did not have any involvement in the writing of the manuscript or the decision to submit it for publication. Additionally, there was partial support from the Oxford NIHR Biomedical Research Centre.

## Footnote

*Reporting Checklist:* The authors have completed the STROBE reporting checklist. Available at <https://tldr.amegroups.com/article/view/10.21037/tlcr-24-241/rc>

*Data Sharing Statement:* Available at <https://tldr.amegroups.com/article/view/10.21037/tlcr-24-241/dss>

[com/article/view/10.21037/tlcr-24-241/dss](https://tldr.amegroups.com/article/view/10.21037/tlcr-24-241/dss)

*Peer Review File:* Available at <https://tldr.amegroups.com/article/view/10.21037/tlcr-24-241/prf>

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at <https://tldr.amegroups.com/article/view/10.21037/tlcr-24-241/coif>). N.L.B. receives grants/consultancy fees/payments from Theramex, F. Hoffmann-La Roche, Innovate UK, and Sleep Universal Limited. A.G. is the Vice President, and Global Head of Data Sciences, at Odysseus Data Services Inc, since September 2021. D.P.A. research group has received research grants from Amgen, Chiesi-Taylor, Lilly, Janssen, Novartis, and from UCB Biopharma; consultancy fees (paid to his department) from Astra Zeneca and UCB Biopharma; and other financial or non-financial interests from Amgen, Astellas, Janssen, Synapse Management Partners and UCB Biopharma have funded or supported training programmes organised by D.P.A. and his department. The other authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The use of CPRD data in this study was approved by the Independent Scientific Advisory Committee (No. 22\_001843). The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Individual consent for this retrospective analysis was waived.

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021;71:209-49.

2. Leiter A, Veluswamy RR, Wisnivesky JP. The global burden of lung cancer: current status and future trends. *Nat Rev Clin Oncol* 2023;20:624-39.
3. James J. Smoking, information, and education: The Royal College of Physicians and the new public health movement. *Journal of Policy Analysis and Management* 2024;43:446-71.
4. Ang L, Chan CPY, Yau WP, et al. Association between family history of lung cancer and lung cancer risk: a systematic review and meta-analysis. *Lung Cancer* 2020;148:129-37.
5. Palma-Lara I, Martínez-Castillo M, Quintana-Pérez JC, et al. Arsenic exposure: A public health problem leading to several cancers. *Regul Toxicol Pharmacol* 2020;110:104539.
6. Lorenzo-González M, Torres-Durán M, Barbosa-Lorenzo R, et al. Radon exposure: a major cause of lung cancer. *Expert Rev Respir Med* 2019;13:839-50.
7. Kwak K, Kang D, Paek D. Environmental exposure to asbestos and the risk of lung cancer: a systematic review and meta-analysis. *Occup Environ Med* 2022;79:207-14.
8. Office for National Statistics [Internet]. 2021. Smoking prevalence in the UK and the impact of data collection changes: 2020. Accessed 26 Feb 2024. Available online: <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/drugusealcoholandsmoking/bulletins/smokingprevalenceintheukandtheimpactofdatacollectionchanges/2020>
9. Martins BNFL, Normando AGC, Rodrigues-Fernandes CI, et al. Global frequency and epidemiological profile of electronic cigarette users: a systematic review. *Oral Surg Oral Med Oral Pathol Oral Radiol* 2022;134:548-61.
10. National Health Service, 2022. Accessed 26 Feb 2024. Available online: <https://digital.nhs.uk/news/2022/decrease-in-smoking-and-drug-use-among-school-children-but-increase-in-vaping-new-report-shows#:~:text=The%20number%20of%20young%20people,1%2C%20-statistics%20published%20today%20show>
11. Kowalczyk A, Jassem J. Multidisciplinary team care in advanced lung cancer. *Transl Lung Cancer Res* 2020;9:1690-8.
12. Michelotti A, de Scordilli M, Bertoli E, et al. NSCLC as the Paradigm of Precision Medicine at Its Finest: The Rise of New Druggable Molecular Targets for Advanced Disease. *Int J Mol Sci* 2022;23:6748.
13. Cancer Research UK, [Internet]. 2024. Lung cancer mortality statistics. Accessed 26 Feb 2024. Available online: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer/mortality#heading=Two>
14. UK Government, June 2023, Accessed 26 Feb 2024. Available online: <https://www.gov.uk/government/news/new-lung-cancer-screening-roll-out-to-detect-cancer-sooner>
15. O'Dowd EL, Lee RW, Akram AR, et al. Defining the road map to a UK national lung cancer screening programme. *Lancet Oncol* 2023;24:e207-18.
16. Herrett E, Gallagher AM, Bhaskaran K, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* 2015;44:827-36.
17. Hripcsak G, Duke JD, Shah NH, et al. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. *Stud Health Technol Inform* 2015;216:574-8.
18. Voss EA, Makadia R, Matcho A, et al. Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. *J Am Med Inform Assoc* 2015;22:553-64.
19. Gilbert J, Rao G, Schuemie M, et al. Cohort diagnostics: diagnostics for OHDSI studies. 2020. Accessed 3 Jun 2024. Available online: <https://ohdsi.github.io/CohortDiagnostics>
20. Gallagher AM, Dedman D, Padmanabhan S, et al. The accuracy of date of death recording in the Clinical Practice Research Datalink GOLD database in England compared with the Office for National Statistics death registrations. *Pharmacoepidemiol Drug Saf* 2019;28:563-9.
21. Eurostat Taskforce. Revision of the European Standard Population - Report of Eurostat's task force. 2013. Available online: <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-ra-13-028> (accessed 2nd August 2023).
22. Raventós B, Català M, Du M, et al. IncidencePrevalence: An R package to calculate population-level incidence rates and prevalence using the OMOP common data model. *Pharmacoepidemiol Drug Saf* 2024;33:e5717.
23. Therneau TM, Grambsch PM. *Modeling Survival Data: Extending the Cox Model*. Springer, New York. ISBN 0-387-98784-3; 2000.
24. NHS England [Internet]. 2023 [cited 2024 Jan 9]. Cancer Prevalence. National Disease Registration Service. Available online: [https://www.cancerdata.nhs.uk/incidence\\_and\\_mortality](https://www.cancerdata.nhs.uk/incidence_and_mortality)
25. Cancer Research UK. Lung Cancer Statistics. Accessed 26 Feb 2024. Available online: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer/mortality#heading=Two>



- cancer-type/lung-cancer
26. Gysling S, Morgan H, Ifesemen OS, et al. The Impact of COVID-19 on Lung Cancer Incidence in England: Analysis of the National Lung Cancer Audit 2019 and 2020 Rapid Cancer Registration Datasets. *Chest* 2023;163:1599-607.
  27. Iyen-Omofoman B, Hubbard RB, Smith CJ, et al. The distribution of lung cancer across sectors of society in the United Kingdom: a study using national primary care data. *BMC Public Health* 2011;11:857.
  28. Ganti AK, Klein AB, Cotarla I, et al. Update of Incidence, Prevalence, Survival, and Initial Treatment in Patients With Non-Small Cell Lung Cancer in the US. *JAMA Oncol* 2021;7:1824-32.
  29. Gridelli C, Rossi A, Carbone DP, et al. Non-small-cell lung cancer. *Nat Rev Dis Primers* 2015;1:15009.
  30. Guarga L, Amejjide A, Marcos-Gragera R, et al. Trends in lung cancer incidence by age, sex and histology from 2012 to 2025 in Catalonia (Spain). *Sci Rep* 2021;11:23274.
  31. Fu Y, Liu J, Chen Y, et al. Gender disparities in lung cancer incidence in the United States during 2001-2019. *Sci Rep* 2023;13:12581.
  32. DeSantis CE, Miller KD, Dale W, et al. Cancer statistics for adults aged 85 years and older, 2019. *CA Cancer J Clin* 2019;69:452-67.
  33. Royal College of Surgeons of England. National Lung Cancer Audit: State of the Nation Report 2023. Accessed 26 Feb 2024. Available online: <https://www.lungcanceraudit.org.uk/reports-publications/nlca-state-of-the-nation-report-2023-version-3/>
  34. NHS England. National Disease Registration Service. Cancer Survival in England: adult, stage at diagnosis, childhood and geographical patterns. Available online: <https://www.cancerdata.nhs.uk/survival/cancersurvivalengland>
  35. Lu T, Yang X, Huang Y, et al. Trends in the incidence, treatment, and survival of patients with lung cancer in the last four decades. *Cancer Manag Res* 2019;11:943-53.
  36. Allemani C, Matsuda T, Di Carlo V, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet* 2018;391:1023-75.
  37. Francisci S, Minicozzi P, Pierannunzio D, et al. Survival patterns in lung and pleural cancer in Europe 1999-2007: Results from the EURO CARE-5 study. *Eur J Cancer* 2015;51:2242-53.
  38. Smith DR, Behzadnia A, Imawana RA, et al. Exposure-lag response of smoking prevalence on lung cancer incidence using a distributed lag non-linear model. *Sci Rep* 2021;11:14478.
  39. Feliu A, Filippidis FT, Joossens L, et al. Impact of tobacco control policies on smoking prevalence and quit ratios in 27 European Union countries from 2006 to 2014. *Tob Control* 2019;28:101-9.
  40. Jackson SE, Beard E, Angus C, et al. Moderators of changes in smoking, drinking and quitting behaviour associated with the first COVID-19 lockdown in England. *Addiction* 2022;117:772-83.
  41. Tattan-Birch H, Jarvis MJ. Children's exposure to second-hand smoke 10 years on from smoke-free legislation in England: Cotinine data from the Health Survey for England 1998-2018. *Lancet Reg Health Eur* 2022;15:100315.
  42. Bracken-Clarke D, Kapoor D, Baird AM, et al. Vaping and lung cancer - A review of current data and recommendations. *Lung Cancer* 2021;153:11-20.
  43. Stone PW, Osen M, Ellis A, et al. Prevalence of Chronic Obstructive Pulmonary Disease in England from 2000 to 2019. *Int J Chron Obstruct Pulmon Dis* 2023;18:1565-74.
  44. Public Health England. UK National Radon Action Plan. 2018 Dec. Accessed 26 Feb 2024. Available online: <https://www.gov.uk/government/publications/uk-national-radon-action-plan>
  45. Aßenmacher M, Kaiser JC, Zaballa I, et al. Exposure-lag-response associations between lung cancer mortality and radon exposure in German uranium miners. *Radiat Environ Biophys* 2019;58:321-36.
  46. Turner MC, Andersen ZJ, Baccarelli A, et al. Outdoor air pollution and cancer: An overview of the current evidence and public health recommendations. *CA Cancer J Clin* 2020. [Epub ahead of print]. doi: 10.3322/caac.21632.
  47. Huang Y, Zhu M, Ji M, et al. Air Pollution, Genetic Factors, and the Risk of Lung Cancer: A Prospective Study in the UK Biobank. *Am J Respir Crit Care Med* 2021;204:817-25. Erratum in: *Am J Respir Crit Care Med* 2022;205:1254.
  48. May L, Shows K, Nana-Sinkam P, et al. Sex Differences in Lung Cancer. *Cancers (Basel)* 2023;15:3111.
  49. Ragavan M, Patel MI. The evolving landscape of sex-based differences in lung cancer: a distinct disease in women. *Eur Respir Rev* 2022;31:210100.
  50. Chien LH, Jiang HF, Tsai FY, et al. Incidence of Lung Adenocarcinoma by Age, Sex, and Smoking Status in Taiwan. *JAMA Netw Open* 2023;6:e2340704.

51. Xiong WM, Xu QP, Li X, et al. The association between human papillomavirus infection and lung cancer: a system review and meta-analysis. *Oncotarget* 2017;8:96419-32.
52. Barrera-Rodriguez R, Morales-Fuentes J. Lung cancer in women. *Lung Cancer (Auckl)* 2012;3:79-89.
53. Mollerup S, Berge G, Baera R, et al. Sex differences in risk of lung cancer: Expression of genes in the PAH bioactivation pathway in relation to smoking and bulky DNA adducts. *Int J Cancer* 2006;119:741-4.
54. Stapelfeld C, Dammann C, Maser E. Sex-specificity in lung cancer risk. *Int J Cancer* 2020;146:2376-82.
55. Denton E, Conron M. Improving outcomes in lung cancer: the value of the multidisciplinary health care team. *J Multidiscip Healthc* 2016;9:137-44.
56. Pillay B, Wootten AC, Crowe H, et al. The impact of multidisciplinary team meetings on patient assessment, management and outcomes in oncology settings: A systematic review of the literature. *Cancer Treat Rev* 2016;42:56-72.
57. Navani N, Baldwin DR, Edwards JG, et al. Lung Cancer in the United Kingdom. *J Thorac Oncol* 2022;17:186-93.
58. Cai H, Wang Y, Qin D, et al. Advanced surgical technologies for lung cancer treatment: Current status and perspectives. *Engineered Regeneration* 2023;4:55-67.
59. Ironmonger L, Ohuma E, Ormiston-Smith N, et al. An evaluation of the impact of large-scale interventions to raise public awareness of a lung cancer symptom. *Br J Cancer* 2015;112:207-16.
60. Ball S, Hyde C, Hamilton W, et al. An evaluation of a national mass media campaign to raise public awareness of possible lung cancer symptoms in England in 2016 and 2017. *Br J Cancer* 2022;126:187-95.
61. Yang SC, Wang JD, Wang SY. Considering lead-time bias in evaluating the effectiveness of lung cancer screening with real-world data. *Sci Rep* 2021;11:12180.
62. Roland M. Linking physicians' pay to the quality of care--a major experiment in the United kingdom. *N Engl J Med* 2004;351:1448-54.
63. Swerdlow AJ. Cancer Registration in England and Wales: Some Aspects Relevant to Interpretation of the Data. *J R Stat Soc* 1986;149:146-60.
64. Sarfati D, Blakely T, Pearce N. Measuring cancer survival in populations: relative survival vs cancer-specific survival. *Int J Epidemiol* 2010;39:598-610.
65. Arhi CS, Bottle A, Burns EM, et al. Comparison of cancer diagnosis recording between the Clinical Practice Research Datalink, Cancer Registry and Hospital Episodes Statistics. *Cancer Epidemiol* 2018;57:148-57.
66. Strongman H, Williams R, Bhaskaran K. What are the implications of using individual and combined sources of routinely collected data to identify and characterise incident site-specific cancers? a concordance and validation study using linked English electronic health records data. *BMJ Open* 2020;10:e037719.
67. Mincuzzi A, Carone S, Galluzzo C, et al. Gender differences, environmental pressures, tumor characteristics, and death rate in a lung cancer cohort: a seven-years Bayesian survival analysis using cancer registry data from a contaminated area in Italy. *Front Public Health* 2023;11:1278416.
68. Rudin CM, Brambilla E, Faivre-Finn C, et al. Small-cell lung cancer. *Nat Rev Dis Primers* 2021;7:3.

**Cite this article as:** Corby G, Barclay NL, Tan EH, Burn E, Delmestri A, Duarte-Salles T, Golozar A, Man WY, Tietzova I, Prieto-Alhambra D, Newby D. Incidence, prevalence, and survival of lung cancer in the United Kingdom from 2000–2021: a population-based cohort study. *Transl Lung Cancer Res* 2024;13(9):2187-2201. doi: 10.21037/tlcr-24-241