Ethnic disparities in lung cancer incidence and differences in diagnostic characteristics: a population-based cohort study in England

Daniel Tzu-Hsuan Chen,[a] Jennifer Hirst,[a] Carol A. C. Coupland,[b] Weiqi Liao,[c] David R. Baldwin,[d] and Julia Hippisley-Cox[a,*]

[a]Nuffield Department of Primary Health Care Sciences, University of Oxford, UK
[b]Centre for Academic Primary Care, School of Medicine, University of Nottingham, Nottingham, UK
[c]Department of Cardiovascular Sciences, University of Leicester, Leicester, UK
[d]Department of Respiratory Medicine, Nottingham University Hospitals and University of Nottingham, Nottingham, UK

The Lancet Regional Health - Europe 2025;48: 101124

Published Online xxx https://doi.org/10.1016/j.lanepe.2024.101124

Summary

Background Lung cancer is a leading cause of mortality, yet disparities in lung cancer across different sociodemographic groups in the UK remain unclear. This study investigates ethnicity and sociodemographic disparities and differences in lung cancer in a nationally representative English cohort, aiming to highlight inequalities and promote equitable access to diagnostic advancements.

Methods We conducted a population-based cohort study using health care records from QResearch, a large primary care database in England. The study included adults aged 25 and over, spanning the period of 2005–2019. Lung cancer incidence rates were calculated using age-standardized methods. Multinomial logistic regression was applied to assess associations between ethnicity/sociodemographic factors and diagnostic characteristics (histological type, stage, and cancer grade), adjusting for confounders.

Findings From a cohort of over 17.5 million people, we identified disparities in incidence rates across ethnic groups from 2005 to 2019. Analysis of 84,253 lung cancer cases revealed that younger woman and Individuals of Indian, other Asian, Black African, Caribbean and Chinese backgrounds had a significantly higher risks of adenocarcinoma compared with squamous cell carcinoma than their White counterparts (relative risk ratios [RRR] spanning from 1.52 (95% CI 1.18–1.94) to 2.69 (95% CI 1.43–5.05). Men and current smokers were more likely to be diagnosed at an advanced stage than women and never smokers (RRR: 1.72 [95% CI 1.56–1.90]–2.45 [95% CI 2.16–2.78]). Socioeconomic deprivation was associated with higher risks of moderate or poorly differentiated adenocarcinoma compared with well differentiated (RRRs between 1.35 [CI: 1.02–1.79] and 1.37 [1.05–1.80]).

Interpretation Our study highlights significant differences in lung cancer incidence and in lung cancer diagnostic characteristics related to ethnicity, deprivation and other demographic factors. These findings have important implications for the provision of equitable screening and prevention programmes to mitigate health inequalities.

Funding DART (The Integration and Analysis of Data using Artificial Intelligence to Improve Patient Outcomes with Thoracic Diseases) project, Innovate UK (UK Research and Innovation), QResearch® and grants from the NIHR Biomedical Research Centre (Oxford), John Fell Oxford University Press Research Fund, Cancer Research UK, and the Oxford Wellcome Institutional Strategic Support Fund.

Keywords: Lung cancer; Cancer epidemiology; Health disparities; Cancer prevention; Ethnic disparities; Cohort study

Introduction

Lung cancer is a leading cause of global mortality, responsible for over 1.8 million deaths annually and has a poor prognosis.[1] The United Kingdom (UK) also faces a significant burden, with lung cancer causing more than 35,000 deaths each year.[2]

Early detection and diagnosis are essential for improving lung cancer outcomes. A key advancement in

*Corresponding author. Nuffield Department of Primary Care Health Sciences, Radcliffe Observatory Quarter, Woodstock Road, Oxford, UK.
E-mail address: julia.hippisley-cox@phc.ox.ac.uk (J. Hippisley-Cox).

Research in context

Evidence before this study

We searched for epidemiological studies investigating the association of ethnic and/or sociodemographic disparities with lung cancer outcomes, published up until July 2024 (with no specified earliest date), in PubMed, Scopus, and Web of Science using relevant terms ('lung cancer' AND ('ethnic disparities' OR 'socioeconomic status' OR 'smoking' OR 'cancer diagnostic characteristics')). We included studies focusing on ethnic and sociodemographic factors related to lung cancer incidence, diagnosis, and survival. Our search found several US studies addressing lung cancer disparities and smoking behaviours, demonstrating the importance of these factors in diagnosis and survival outcomes. However, only a few studies were UK-based, and those often lacked sufficient representativeness or failed to account for key factors like socioeconomic status and smoking. Additionally, while low-dose computerized tomography (LDCT) screening has proven successful in reducing lung cancer mortality, these interventions risk exacerbating health inequalities, as suggested by the inverse equity hypothesis. This study seeks to fill these gaps by quantifying the impact of ethnicity, socioeconomic status, and smoking behaviours on lung cancer outcomes within a nationally representative cohort, emphasizing the need to address these disparities in future health strategies.

Added value of this study

This study, spanning from 2005 to 2019, stands as one of the largest and most comprehensive analyses of ethnic disparities in lung cancer within an English primary care cohort, revealing distinct trends in lung cancer incidence among ethnic groups over time. Our findings show variations in histological cancer types between ethnic groups, despite adjustment for smoking, add novel and strong evidence to the existing body of knowledge suggesting that these differences might not only be attributed to smoking habits but could also be influenced by factors such as ethnicity/genetic predispositions and differences in sociodemographic backgrounds. These findings, coupled with notable gender and racial differences in different cancer diagnostics, underscore the importance of taking into account the unique ethnic and socioeconomic factors of different ethnic groups.

Implications of all the available evidence

The findings suggested that tailored lung cancer prevention programmes should take into account the diverse characteristics of ethnic groups and incorporate ethnicity-specific factors into clinical practices to address health inequalities. While our study does not examine screening or treatment access directly, these insights are valuable for healthcare providers and policymakers in creating inclusive health policies. Such policies could contribute to fair access to future screening programs and treatments, potentially reducing health disparities among ethnic minorities and socioeconomically disadvantaged groups.

this area is the adoption of low-dose computerized tomography (LDCT) for lung cancer screening, which has significantly helped reduced mortality rates.[3,4] Many countries, including the United States (US), are now implementing lung cancer screening on a national or regional basis.[1] In June 2022, the UK National Screening Committee recommended targeted LDCT screening for high-risk individuals aged 55–74[5] and focused on geographic areas with high lung cancer incidence and mortality rates in the initial phases of the programme.

Despite the advancements in screening and diagnosis, existing health disparities, as highlighted by the inverse equity hypothesis, suggest that new health interventions often disproportionately benefit those from more affluent backgrounds instead of those who are most in need.[6] Therefore, such preventative measures necessitate careful planning and monitoring to ensure equitable access and implementation for all population segments.

Significant disparities in lung cancer incidence, diagnosis, and survival are evident across various demographics, notably among ethnic minorities.[7–11] These disparities are further compounded by the relationship between socioeconomic status and smoking behaviour, which is a primary risk factor for lung cancer.[12] While extensive research in the United States have shed light on ethnic disparities in lung cancer incidence and outcomes,[13] the UK has lacked comprehensive data that fully explores the impact of ethnicity, age and sex on lung cancer disparities. Previous localised studies within the UK have showed significant differences in lung cancer incidence among ethnic groups but were limited by their scale and focus, often overlooking crucial factors such as smoking status and histological type.[14–17]

Given the under-representation of ethnic minorities in existing lung cancer research, this study aims to fill a critical gap by exploring the epidemiology of lung cancer across diverse ethnic groups within a large, nationally representative cohort from England. By examining disparities in lung cancer incidence over time and diagnostic characteristics, such as stage, histological type, and cancer grade at diagnosis, this study seeks to provide insights that could guide the development of more inclusive health policies and interventions, ensuring equitable access to advancements in cancer detection and treatment.

Methods

Data source, study population and study design

We undertook a population-based cohort study using routinely collected electronic health records (EHRs) from the QResearch database (version 45). QResearch is a large consolidated database with anonymised EHRs of over 35 million patients from 1800+ general practices and is one of the largest health-care databases in England, and a Trusted Research Environment accredited by Health Data Research UK.

We included an open cohort of adult patients aged 25–100 years at cohort entry who were registered with general practices in England and who contributed to the QResearch database from January 1, 2005, the start of the study period and followed up to December 31, 2019, prior to the start of the COVID-19 pandemic in 2020. We excluded patients with pre-existing lung cancer before cohort entry. The broad age range covers the majority of the adult primary care population and provides flexibility to examine lung cancer incidence and select people from different sociodemographic backgrounds to perform subgroup analyses including those who do not currently fall into the age bands considered for screening.

For the analysis of lung cancer diagnostic characteristics, we included only those with an established link to the cancer registry record available up to 2018 and restricted cases to those diagnosed between 2005 and 2018. This was necessary because the cancer registry provides essential details on cancer histology, stage, and grade, which are critical for our analysis. As such, our analysis is based on the individuals with established data linkages.

Study outcomes

The study outcome included incident diagnosis of lung cancer during follow-up to December 31, 2019, recorded on one or more of the four linked data sources in QResearch—primary care and secondary care (i.e., hospital episode statistics [HES]) databases, cancer registry (previously from Public Health England, now part of NHS England), and death registry (from the Office for National Statistics [ONS]). We used the earliest date recorded on any of the four data sources as the date of lung cancer diagnosis to ensure comprehensive capture of new lung cancer cases across the various data sources.

We used SNOMED-CT codes to identify lung cancer diagnoses from the general practice record, and ICD-10 codes to identify events from HES, and cancer and death registries. The primary outcome was from one or more of the linked data sources: HES, and ONS.

We also considered three major cancer diagnosis characteristics of lung cancer, including histological subtypes of lung cancer, stage at diagnosis, and histological grading of lung adenocarcinoma to examine ethnic and sociodemographic disparities and differences in lung cancer diagnostic factors.

Definitions of cancer histology[18] were based on the ICD-10-O-2 system, and were defined as adenocarcinoma, squamous cell, small cell and large cell carcinoma, others and unspecified in the study. Cancer staging was based on the TNM Classification of Malignant Tumours System (Stages I to IV, and not recorded). The World Health Organization (WHO) grading system classifies lung adenocarcinomas based on tumour cell differentiation under microscopic examinations.[19] This approach guides the reported results of our lung adenocarcinoma grading analysis, defined as well, moderately and poorly differentiated, undifferentiated and not recorded according to patients' records.

Study exposures and variables

Our main exposures of interest were self-assigned ethnic group as recorded in primary care records (White, Indian, Pakistani, Bangladeshi, Chinese, Other Asian, Caribbean, Black African, Other and not recorded) and quintile of Townsend deprivation score (Q1–Q5 where Q1 is most affluent and Q5 is most deprived).[20] Secondary exposures included age at lung cancer diagnosis, sex and most recent smoking status before lung cancer diagnosis (for descriptive analysis, we used five groups: never, former, light [1–9 cigarettes per day], moderate [10–19 cigarettes per day] and heavy smokers [≥20 cigarettes per day]; for the regression models, we simplified this to three categories: never, former, and current smokers).

Additional variables included were body mass index (BMI) recorded closest to diagnosis (below 18.5, 18.5–24.9, 25–29.9, 30–34.9, 35 and above, missing), broad geographical regions in England (North, Midlands, Southern England and London), and lung related co-morbidities (tuberculosis, chronic obstructive pulmonary disease [COPD] and asthma).

Statistical analysis

We report descriptive statistics to characterise the demographic and clinical features of the study cohort and those diagnosed with lung cancer during the study period. We calculated crude and European age-standardised[21] annual lung cancer incidence rates, expressed per 100,000 person-years, and plotted these by ethnicity, sex, age groups, socioeconomic deprivation (measured by Townsend quintiles), and smoking status. Incidence rates were calculated by dividing the number of new lung cancer cases diagnosed between 2005 and 2019 by the total person-years at risk. Person-years were summed from each individual's entry into the study until the earliest of lung cancer diagnosis, death, loss to follow-up, or the study's end on December 31, 2019. Age standardisation used the direct method with the European Standard Population as the reference to adjust for age distribution differences across subgroups.

We used adjusted multinomial logistic regression models in those with a diagnosis of lung cancer between

2005 and 2018 to investigate ethnic and sociodemographic differences (ie. Townsend quintiles, sex and age groups) associated with (1) different histological subtypes of lung cancer, with squamous cell carcinoma as the reference group, (2) stage at diagnosis with stage I as the reference group, and (3) histological grading among lung adenocarcinoma with well-differentiated as the reference group. The results report the relative risk ratios (RRRs) obtained from the multinomial logistic regression analysis to investigate the association between the study exposures and lung cancer outcomes. The RRR is defined as a measure that compares the relative likelihood of one outcome category relative to a reference outcome category between groups, while holding other covariates constant.

All the regression analyses were adjusted for BMI, smoking status, geographical regions, and lung related co-morbidities. For categorical variables with missing data, we categorised them as "not recorded" to ensure all available data was used in the analysis. This approach allowed us to conduct analyses without excluding any records, ensuring that the dataset remained as comprehensive as possible.

We also conducted sensitivity analyses using ordinal regression models to examine differences in histological grading and stage at diagnosis, and mixed effect multilevel multinomial logistic models to account for clustering within GP practices. Furthermore, we included three additional two-way interaction terms in the regression models to examine whether the associations between ethnicity (categorised into broader groups of White, Asian [Indian, Pakistani, Bangladeshi, Other Asian], Black [Caribbean, Black African], Chinese and Others[22] for examining interactions) varied by patients' sex, deprivation quintile, and smoking status in terms of lung cancer diagnostic characteristics. The statistical significance level was set as 0.05 in this study. Data were managed and analysed using Stata 18.0. The QResearch ethics approval by the East Midlands-Derby Research Ethics Committee [reference 18/EM/0400].

Role of the funding source

The funding source supported the development of technical and governance infrastructure to enable data access for research, as well as the researchers' time to conduct the study. The funders had no role in the study design, protocol development, data curation, analysis, interpretation, or manuscript writing.

Results

Characteristics of the study cohort

There were 17,990,154 individuals in the initial primary care cohort, with 17,587,168 remaining after excluding those with prior lung cancer and those who entered after December 31, 2019. We used this cohort, to calculate lung cancer incidence rates from 2005 to 2019 by ethnicity, sex, age, deprivation, and smoking status. Of these 84,253 were diagnosed between 2005 and 2018 and included in analyses of lung cancer characteristics (Table 1).

Among the 84,253 lung cancer cases, ethnicity was recorded for 57,249 (67.9%) individuals, while it was not recorded for the remaining 32.1%. Of those with recorded ethnicity, the vast majority (95.3%) were White, accounting for 64.8% (n = 54,563) of the total lung cancer cases, followed by smaller proportions of Caribbean or Black African (0.7% and 0.3%, respectively), Indian (0.5%), Bangladeshi (0.4%), Pakistani (0.3%), other Asian or Chinese (0.3%, and 0.2%, respectively), and other ethnicities (0.6%). The majority of lung cancer cases were male (55.0%). The proportion of lung cancer cases that were male was higher among patients of Pakistani (75.4%), Bangladeshi (78.3%), Caribbean (71.6%), Indian (66.4%), and other Asian (63.6%) backgrounds.

Those with a lung cancer diagnosis from non-White ethnic groups largely came from more deprived backgrounds (Townsend quintiles Q4–Q5), particularly among patients from Bangladeshi, Caribbean and black African backgrounds.

Indian, Black African, Chinese, and other Asian individuals showed a higher prevalence of adenocarcinoma diagnoses (41.1%–53.1%) than other ethnic groups, often with poorly differentiated tumours (16.0%–23.3%). Higher proportions of people from Chinese, other Asian, Caribbean, and Black African backgrounds were diagnosed at stage 4 (43.8%–48.6%) compared with 34.0%–38.0% in the other groups, showing distinct patterns of lung cancer manifestation in different populations.

Trends in incidence of lung cancer by ethnicity

During the study period, the age-standardised incidence rates of lung cancer varied between the different ethnic groups, with men showing higher rates than women in each group. The trends in the age-standardized incidence rates of lung cancer by ethnicity between 2005 and 2019, are detailed in Fig. 1 and Supplementary Table S1.

Over the most recent decade (2010–2019), Bangladeshi men exhibited the highest average age-standardized incidence rate at 170.2 per 100,000 person-years, followed by White (138.6), Chinese (115.1) and Caribbean (114.6) men. The lowest rates were observed in Black African (60.4) and Indian men (50.3). During the same period, Chinese men experienced the largest decline in rates, dropping 39.2% from 163.6 (95% confidence interval [CI]: 40.4–286.9) to 99.4 (CI: 37.3–161.5) per 100,000 person-years. White and Bangladeshi men also saw declines of 20% (from 145.9 [CI: 140.4–151.5] to 115.4 [CI: 111.1–119.7] per 100,000 person-years) and 18.3% (from 172.4 [CI: 96.7–248.0] to 140.8 [CI: 72.1.5–209.4] per 100,000 person-years),

	White	Indian	Pakistani	Bangladeshi	Other Asian	Caribbean	Black African	Chinese	Other	Not recorded	Total
Total %, (N)	64.8 (54,563)	0.5 (384)	0.3 (256)	0.4 (336)	0.3 (231)	0.7 (599)	0.3 (212)	0.2 (128)	0.6 (540)	32.1 (27,004)	84,253
Sex											
Female	45.9 (25,043)	33.6 (129)	24.6 (63)	21.7 (73)	36.4 (84)	28.4 (170)	38.2 (81)	47.7 (61)	39.3 (212)	44.5 (12,029)	45.0 (37,945)
Male	54.1 (29,520)	66.4 (255)	75.4 (193)	78.3 (263)	63.6 (147)	71.6 (429)	61.8 (131)	52.3 (67)	60.7 (328)	55.5 (14,975)	55.0 (46,308)
Age at diagnosis											
<45	0.8 (427)	2.6 (10)	3.9 (10)	3.0 (10)	3.0 (7)	1.8 (11)	8.0 (17)	3.9 (5)	7.0 (38)	0.9 (231)	0.9 (766)
45–54	4.5 (2447)	6.3 (24)	8.6 (22)	7.7 (26)	12.1 (28)	9.2 (55)	23.1 (49)	10.2 (13)	14.6 (79)	4.8 (1299)	4.8 (4042)
55–64	16.0 (8750)	20.8 (80)	22.3 (57)	14.0 (47)	24.2 (47)	18.9 (113)	24.5 (52)	21.1 (27)	23.5 (127)	16.1 (4.337)	16.2 (13,646)
65–74	32.9 (17,924)	31.3 (120)	24.2 (62)	32.1 (108)	30.3 (108)	27.2 (163)	24.5 (52)	23.4 (30)	23.2 (125)	29.7 (8020)	31.7 (26,674)
75+	45.9 (25,015)	39.1 (150)	41.0 (105)	43.2 (145)	30.0 (145)	42.9 (257)	19.8 (42)	41.4 (53)	31.7 (171)	48.6 (13,117)	46.4 (39,125)
Townsend quintile											
Q1 (least deprived)	20.6 (11,255)	7.8 (30)	3.1 (8)	1.5 (5)	9.5 (22)	1.5 (9)	2.4 (5)	13.3 (17)	7.2 (39)	25.1 (6785)	21.6 (18,175)
Q2	21.0 (11,443)	17.4 (67)	12.1 (31)	3.0 (10)	13.4 (31)	3.7 (22)	5.2 (11)	16.4 (21)	8.1 (44)	24.4 (6598)	21.7 (18,278)
Q3	21.4 (11,698)	24.0 (92)	25.8 (66)	4.2 (14)	22.1 (51)	11.9 (71)	12.3 (26)	24.2 (31)	14.4 (78)	21.7 (5872)	21.4 (17,999)
Q4	19.8 (10,826)	29.7 (114)	40.2 (103)	21.1 (71)	28.6 (66)	25.2 (151)	19.8 (42)	16.4 (21)	23.1 (125)	17.6 (4741)	19.3 (16,260)
Q5 (most deprived)	17.0 (9298)	20.8 (80)	18.8 (48)	70.2 (236)	26.4 (61)	57.6 (345)	60.4 (128)	29.7 (38)	47.0 (254)	11.0 (2975)	16.0 (13,463)
Not recorded	0.1 (43)	0.3 (a)	0 (0)	0 (0)	0 (0)	0.2 (a)	0 (0)	0 (0)	0 (0)	0.1 (33)	0.1 (78)
Geographical region[b]											
Midlands	15.1 (8250)	17.2 (66)	17.2 (44)	6.3 (21)	7.4 (17)	16.4 (98)	2.4 (a)	14.8 (19)	10.4 (56)	15.2 (4092)	15.0 (12,668)
Southern England	33.0 (17,993)	10.9 (42)	18.4 (47)	5.4 (18)	23.4 (54)	8.2 (49)	9.4 (20)	16.4 (21)	18.7 (101)	41.8 (11,276)	35.2 (29,621)
Northern England	36.1 (19,712)	8.9 (34)	28.9 (74)	6.0 (20)	7.4 (17)	4.0 (24)	3.8 (8)	19.5 (25)	13.5 (73)	35.9 (9705)	35.2 (29,692)
London	15.8 (8608)	63.0 (242)	35.5 (91)	82.4 (277)	61.9 (143)	71.5 (428)	84.4 (179)	49.2 (63)	57.4 (310)	7.2 (1931)	14.6 (12,272)
BMI category											
Below 18.5 kg/m^2	5.2 (2854)	4.9 (19)	3.9 (10)	4.2 (14)	6.5 (15)	6.5 (39)	3.3 (7)	9.4 (12)	5.7 (31)	4.4 (1185)	5.0 (4186)
18.5–24.9	37.3 (20,352)	45.3 (174)	33.6 (86)	53.0 (178)	45.0 (104)	38.2 (229)	31.6 (67)	54.7 (70)	38.9 (210)	32.4 (8757)	35.9 (30,227)
25–29.9	32.2 (17,592)	33.1 (127)	41.8 (107)	32.1 (108)	29.9 (69)	32.1 (192)	30.2 (64)	26.6 (34)	27.8 (150)	24.9 (6712)	29.9 (25,155)
30–34.9	13.9 (7570)	11.5 (44)	11.7 (30)	6.8 (23)	8.7 (20)	13.5 (81)	18.9 (40)	3.9 (5)	10.9 (59)	9.6 (2590)	12.4 (10,462)
35 and above	5.8 (3185)	2.6 (10)	6.3 (16)	1.2 (a)	3.0 (7)	2.8 (17)	11.3 (24)	0.0 (0)	6.7 (36)	3.7 (989)	5.1 (4288)
Missing	5.5 (3010)	2.6 (10)	2.7 (7)	2.7 (9)	6.9 (16)	6.8 (41)	4.7 (10)	5.5 (7)	10.0 (54)	25.1 (6771)	11.8 (9935)
Smoking category											
Never smoker	6.7 (3634)	38.3 (147)	31.6 (81)	11.3 (38)	29.9 (69)	14.9 (89)	39.2 (83)	44.5 (57)	18.5 (100)	8.3 (2250)	7.8 (6548)
Former smoker	54.8 (29,914)	35.9 (138)	40.6 (104)	54.2 (182)	43.3 (100)	41.9 (251)	33.5 (71)	32.8 (42)	37.0 (200)	44.6 (12,049)	51.1 (43,051)
Light smoker (1–9/day)	28.0 (15,254)	22.7 (87)	24.6 (63)	28.9 (97)	19.0 (44)	35.4 (212)	22.2 (47)	14.8 (19)	35.7 (193)	25.9 (7006)	27.3 (23,022)
Moderate smoker (10–19/day)	5.6 (3071)	1.8 (7)	2.0 (a)	4.2 (14)	3.5 (8)	4.8 (29)	3.8 (8)	4.7 (6)	6.5 (35)	4.8 (1307)	5.3 (4490)
Heavy smoker (20 + day)	4.5 (2463)	1.3 (a)	1.2 (a)	1.2 (a)	3.0 (7)	2.5 (15)	1.4 (a)	2.3 (a)	1.7 (9)	4.1 (1120)	4.3 (3632)
Not recorded	0.4 (227)	0 (0)	0 (0)	0.3 (a)	1.3 (a)	0.5 (a)	0 (0)	0.8 (a)	0.6 (a)	12.1 (3272)	4.2 (3510)
Histological type											
Adenocarcinoma	26.1 (14,251)	41.1 (158)	31.3 (80)	28.0 (94)	44.6 (103)	36.2 (217)	47.2 (100)	53.1 (68)	40.4 (218)	21.9 (5926)	25.2 (21,215)
Squamous cell	19.3 (10,550)	13.3 (51)	20.3 (52)	26.2 (88)	13.9 (32)	16.5 (99)	11.8 (25)	9.4 (12)	11.1 (60)	16.9 (4570)	18.4 (15,539)
Small cell	10.5 (5705)	5.2 (20)	10.2 (26)	6.8 (23)	6.1 (14)	7.8 (47)	6.6 (14)	5.5 (7)	9.4 (51)	11.0 (2969)	10.5 (8876)
Large cell/Other	4.6 (2533)	6.5 (25)	6.3 (16)	3.9 (13)	7.4 (17)	4.6 (28)	7.5 (16)	7.1 (9)	6.7 (36)	3.7 (1005)	4.4 (3697)
Unspecified	39.4 (21,498)	33.9 (130)	31.6 (81)	35.1 (118)	28.1 (65)	34.7 (208)	26.9 (57)	25.0 (32)	32.2 (174)	46.4 (12,521)	41.4 (34,884)

(Table 1 continues on next page)

	White	Indian	Pakistani	Bangladeshi	Other Asian	Caribbean	Black African	Chinese	Other	Not recorded	Total
(Continued from previous page)											
Cancer stage at diagnosis											
Stage 1	11.7 (6410)	11.7 (45)	10.9 (28)	9.8 (33)	10.0 (23)	9.0 (54)	8.5 (18)	16.4 (21)	10.4 (56)	6.1 (1634)	9.9 (8322)
Stage 2	5.7 (3132)	7.3 (28)	8.6 (22)	5.1 (17)	3.9 (9)	3.7 (22)	3.3 (7)	3.9 (5)	4.8 (26)	3.4 (912)	5.0 (4180)
Stage 3	15.8 (8607)	11.5 (44)	20.7 (53)	21.1 (71)	18.2 (42)	13.7 (82)	11.8 (25)	10.9 (14)	16.5 (89)	10.8 (2913)	14.2 (11,940)
Stage 4	38.0 (20,721)	38.0 (146)	34.0 (87)	37.8 (127)	45.0 (104)	44.4 (266)	48.6 (103)	43.8 (56)	41.9 (226)	28.5 (7696)	35.1 (29,532)
Stage not recorded	28.8 (15,693)	31.5 (121)	25.8 (66)	26.2 (88)	22.9 (53)	29.2 (175)	27.8 (59)	25.0 (32)	26.5 (143)	51.3 (13,849)	35.9 (30,279)
Grade at diagnosis (adenocarcinoma, N = 21,215)											
Well differentiated	4.5 (644)	4.4 (7)	10.0 (8)	6.4 (6)	4.9 (5)	5.1 (11)	5.0 (5)	5.9 (a)	4.1 (9)	4.3 (256)	4.5 (955)
Moderately differentiated	13.6 (1942)	11.4 (18)	15.0 (12)	25.5 (24)	5.8 (6)	10.1 (22)	17.0 (17)	5.9 (a)	11.0 (24)	13.5 (797)	13.5 (2866)
Poorly differentiated	20.2 (2873)	19.0 (30)	20.0 (16)	20.2 (19)	23.3 (24)	22.1 (48)	16.0 (16)	25.0 (17)	23.4 (51)	19.6 (1162)	20.1 (4255)
Undifferentiated	0.3 (43)	0.0 (0)	0.0 (0)	0.0 (0)	1.0 (a)	0.5 (a)	0.0 (0)	0.0 (0)	0.0 (0)	0.2 (14)	0.3 (59)
Not recorded	61.4 (8750)	65.2 (103)	55.0 (44)	47.8 (45)	65.1 (67)	62.2 (135)	62.0 (62)	63.3 (43)	61.5 (134)	62.4 (3697)	61.7 (13,080)
Comorbidities											
COPD	33.7 (18,395)	19.3 (74)	21.1 (54)	37.2 (125)	13.9 (32)	19.0 (114)	6.6 (14)	10.2 (13)	17.8 (96)	20.9 (5648)	29.2 (24,565)
Asthma	15.2 (8280)	18.0 (69)	23.0 (59)	19.3 (65)	15.2 (35)	13.2 (79)	10.8 (23)	7.8 (10)	13.9 (75)	10.0 (2708)	13.5 (11,403)
Tuberculosis	2.1 (1120)	9.1 (35)	10.2 (26)	9.5 (32)	3.0 (7)	2.2 (13)	5.7 (12)	5.5 (7)	3.3 (18)	1.6 (423)	2.0 (1693)

Numbers are column percentages with counts in parentheses, %(N). Counts of <5 have been suppressed for confidentiality reasons. [a]Indicates counts below 5 that are suppressed. COPD, Chronic obstructive pulmonary disease. [b]Midlands: East Midlands, West Midlands; Southern England: East of England, South Central, South East, South West; Northern England: North East, North West, Yorkshire & Humber; and London.

Table 1: **Baseline characteristics of 84,253 patients with an incident diagnosis of lung cancer in the study period by demographic and clinical features.**

respectively. Conversely, Pakistani and Black African men showed increases of 87.0% (from 50.1 [CI: 13.2–87.0] to 93.7 [CI: 54.6–132.8]) and 44.7% (from 39.4 [CI: 2.8–76.0] to 57.0 [CI: 27.4–86.6] per 100,000 person-years), respectively.

In the 10 years from 2010 to 2019, White women had the highest average lung cancer incidence rate at 96.9 per 100,000 person-years, with Chinese women following at 75.0. In contrast, the lowest rates were seen in Indian (27.7), Pakistani (26.4), and Black African women (23.0). During this time, the incidence rates among white women slightly increased by 0.3%, from 90.5 (CI: 86.8–94.7) to 90.8 (CI: 87.0–94.0) per 100,000 person-years. Bangladeshi women showed a 1.0% increase in rates, from 54.1 (CI: 7.4–100.7) to 54.6 (CI: 20.0–89.2) per 100,000 person-years, and Indian women saw a higher rise of 36.2%, from 24.6 (CI: 8.1–41.2) to 33.5 (CI: 19.9–47.0) per 100,000 person-years. In contrast, incidence rates generally declined among women from other ethnic groups over time.

Trends in incidence of lung cancer by sex, age, deprivation, and smoking status

The incidence rates displayed contrasting trends between sexes from 2010 to 2018. During this period, the rates for females increased by 3.1%, rising from 91.2 (CI: 87.8–94.6) in 2010 to 94.0 (CI: 90.9–97.1) per 100,000 person-years in 2018. In contrast, the rates for males decreased by 11.7%, dropping from 143.2 (CI: 138.4–147.9) to 126.4 (CI: 122.5–130.3) per 100,000 person-years. However, in 2019, both sexes experienced an over 10% drop in rates compared to 2018. Fig. 2 and Supplementary Table S2 show the trends in lung cancer incidence by sex from 2005 to 2019.

Between 2010 and 2019, the highest lung cancer incidence rates were seen in individuals aged 75 and older for both sexes. During this period, the incidence rate increased slightly for females, while it decreased for males. Age groups below 54 consistently had lower incidence rates throughout the study period (Fig. 2 and Supplementary Table S4).

Between 2010 and 2019 lung cancer incidence rates demonstrated a clear socioeconomic gradient in both men and women. The most deprived quintile (Q5) experienced the highest rates, averaging 214.5 for men and 147.2 for women per 100,000 person-years, whereas the least deprived quintile (Q1) saw lower rates, with averages of 93.5 for men and 61.5 for women per 100,000 person-years. While incidence rates in men have decreased across different levels of deprivation, rates in women have remained relatively stable (Fig. 2 and Supplementary Table S3).

Furthermore, the results highlight marked differences over time in lung cancer incidence based on smoking status across both genders and between ethnic groups. Current smokers had the highest incidence rates for both men and women, with the highest

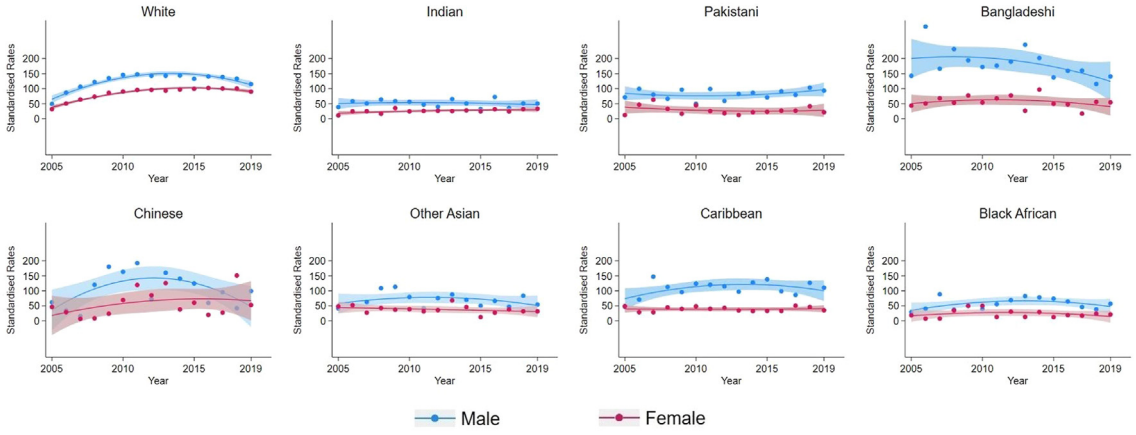

Fig. 1: Age standardised incidence of lung cancer by ethnicity in the primary care population, 2005–2019. Rates per 100,000 person-years adjusted to the standard European population.

observed in heavy smokers—averaging 430.2 for men and 474.3 for women per 100,000 person-years between 2010 and 2019. These rates declined over time. Never smokers showed much lower rates, which also declined throughout the study, particularly in men (from 73.7 in 2005 to 30.1 in 2019 in men; 31.9 to 22.9 in women per 100,000 person-years) (Fig. 2 and Supplementary Table S5). Trends and detailed lung cancer incidence rates by smoking status and ethnicity over time are shown in Supplementary Table S6.

Factors associated with different histological subtypes of lung cancer diagnosis

The multinomial logistic regression analysis in lung cancer cases identified ethnicity, sex, age, and smoking status as significant factors associated with different lung cancer subtypes within the incident lung cancer cases, as detailed in Fig. 3A and Supplementary Table S7.

With squamous cell carcinoma as the reference type, the results showed that individuals of Indian, other Asian, Caribbean, Black African, Chinese, and other ethnic backgrounds had significantly higher risks for adenocarcinoma than their White counterparts, with relative risk ratios (RRR) spanning from 1.52 (95% CI 1.18–1.94) for Caribbean to 2.69 (95% CI 1.43–5.05) for Chinese. Females and younger age groups (particularly <45) had elevated relative risks of developing adenocarcinoma, small cell, and large cell/other lung cancer subtypes in contrast to males and those aged over 75. Current smokers, regardless of their smoking intensity, showed increased risks for small cell lung cancer in

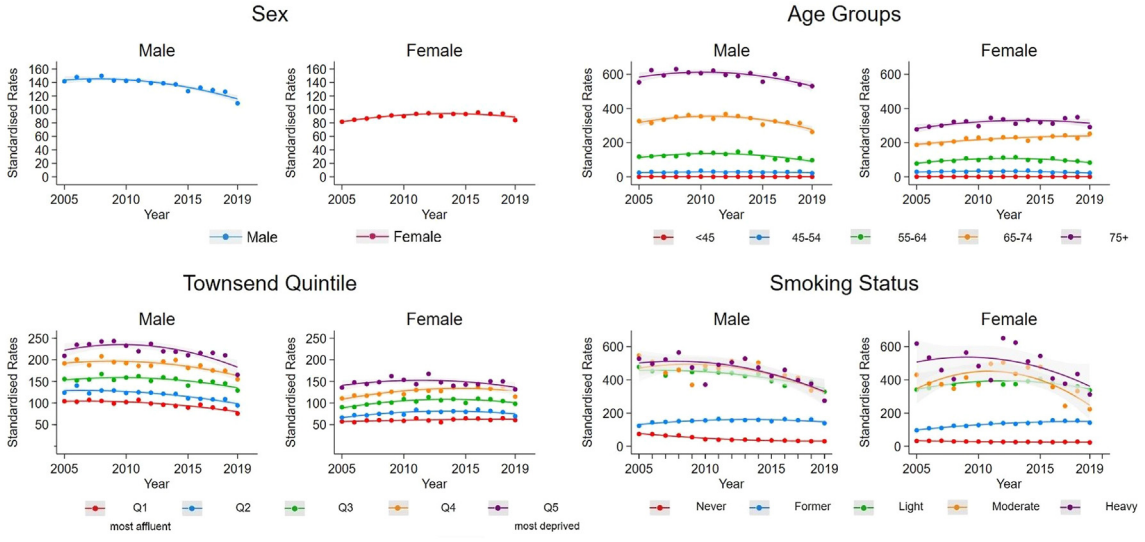

Fig. 2: Age standardised incidence of lung cancer by sex, age groups, Townsend quintiles, and smoking status in the primary care population, 2005–2019. Rates per 100,000 person-years adjusted to the standard European population.

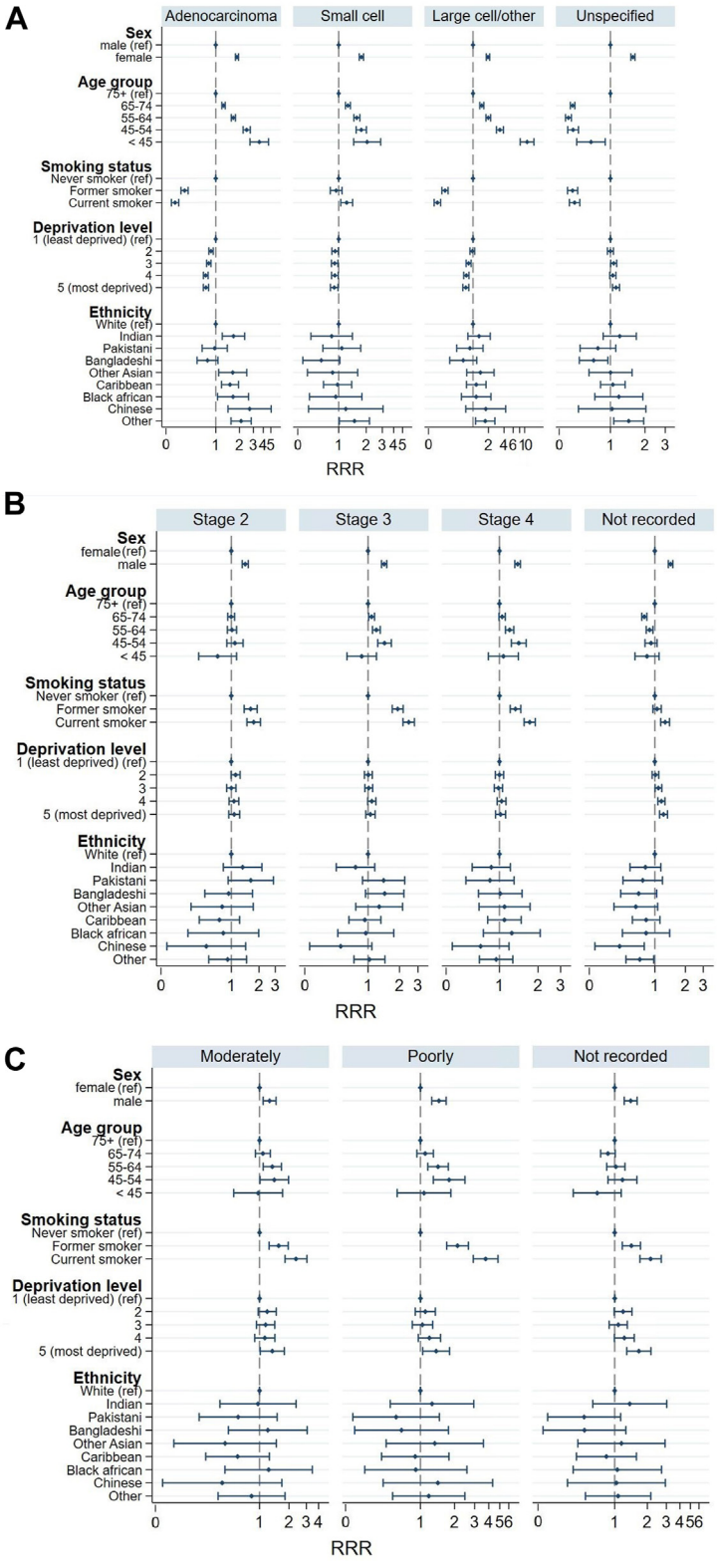

Fig. 3: **A.** Factors significantly associated with an incident diagnosis of different lung cancer subtypes in the primary care population. Squamous cell carcinoma as referent category. The figures show relative risks ratios (RRR) with 95% confidence interval (N = 84,253). Note: association

comparison to squamous cell when compared to never smokers, but decreased risks for adenocarcinoma and large cell/other types. Additionally, analyses of interactions found only significant interactions in former and current smokers whose ethnicity were not recorded. Details of these interactions are provided in Supplementary Table S14.

Factors associated with different stages of lung cancer diagnosis

The modelling results, as shown in Fig. 3B and Supplementary Table S8, identified sex, age, and smoking status as significant factors associated with advanced lung cancer diagnostic stages within the study population. No significant association was found between the stage at diagnosis and levels of socioeconomic deprivation or ethnic backgrounds.

Comparing against stage 1 lung cancer, the analysis showed that men had a higher relative risk of being diagnosed with more advanced stages compared to women, with RRR values of 1.42 (CI: 1.31–1.53) for stage 2, 1.43 (CI: 1.35–1.51) for stage 3 and 1.39 (CI: 1.32–1.46) for stage 4. Individuals in the 45–54 age group had the highest risks of being diagnosed with stages 3 and 4 compared to those above 75, with RRR values of between 1.41 (CI: 1.24–1.61) and 1.44 (CI: 1.24–1.67). Additionally, former and current smokers, were more likely to be diagnosed at advanced stages than never smokers, with RRRs between 1.33 (95% CI: 1.21–1.47) and 2.54 (95% CI 2.11–3.06). However, the only significant interaction terms were in former and current smokers whose ethnicity was not-recorded (Supplementary Table S15).

Factors associated with different grades of lung adenocarcinoma diagnosis

Fig. 3C and Supplementary Table S9 show that sex, age, socioeconomic deprivation, and smoking status are significant factors associated with the grades of diagnoses within the lung adenocarcinoma cases.

With well-differentiated grades as the reference, males and those in the 45–54 and 55–64 age groups had higher risks (RRR between 1.35 [CI: 1.09–1.67] and 1.79 [CI: 1.3–2.47]) of being diagnosed with moderately or poorly differentiated adenocarcinoma. Both former and current smokers had a higher risk of aggressive lung cancer grades (moderately, poorly and undifferentiated) compared never smokers. Moreover, individuals from the most deprived backgrounds (Q5) were at a greater risk of being diagnosed with moderately (RRR: 1.35 [CI: 1.02–1.79]) or poorly differentiated (RRR: 1.37 [1.05–1.8]) lung adenocarcinoma compared to those from the most affluent background. Interaction analysis showed no significant interactions between ethnicity and sex, smoking status and deprivation quintile. Results from all sensitivity analyses are in Supplementary Tables S10–S13, with interaction analysis details in Supplementary Tables S14 and S15.

Discussion

Our population-based study, which is the largest study to date, has identified varying trends in lung cancer incidence rates highlighting some disparities among non-white ethnic groups, reflecting the diverse nature of lung cancer epidemiology across ethnic backgrounds. This indicates the need for ethnic-specific prevention and equitable healthcare access that are sensitive to these ethnic groups. We also found substantial differences in lung cancer incidence rates between different smoking statuses and men and women within each ethnic group, indicating importance of intersectionality. Notably, among Bangladeshi, Chinese, Pakistani and Caribbean populations, these differences are largely due to the varying prevalence of smoking between males and females, particularly in Asian and Southeast Asian communities.[1,23] Over the past few decades, smoking prevalence has generally declined across most ethnic groups in the UK, but these trends have not been uniform. Smoking prevalence among South Asian men remain relatively high compared to other ethnic groups,

estimates (relative risk ratios) were obtained from multinomial logistic regression models, with squamous cell carcinoma as referent group and all listed variables in the models and controlled for BMI, geographical regions in England and lung related co-morbidities (tuberculosis, COPD and asthma). The referent category of the covariates is indicated as a diamond symbol on the first row of each covariate. Age group refers to age at diagnosis. The X axis is displayed on a logarithmic scale. RRRs for ethnicity in the not recorded category are omitted, and can be found in Supplementary Table S7. **B.** Factors significantly associated with an incident diagnosis of different lung cancer diagnostic stages in the primary care population. Stage 1 as referent category. The figures show relative risks ratios (RRR) with 95% confidence interval (N = 84,253). Note: association estimates (relative risk ratios) were obtained from multinomial logistic regression models, with stage 1 as referent group and all listed variables in the models and controlled for BMI, geographical regions in England and lung related co-morbidities (tuberculosis, COPD and asthma). The referent category of the covariates is indicated as a diamond symbol on the first row of each covariate. Age group refers to age at diagnosis. The X axis is displayed on a logarithmic scale. RRRs for ethnicity in the not recorded category are omitted, and can be found in Supplementary Table S8. **C.** Factors significantly associated with an incident diagnosis of different diagnostic grades for lung adenocarcinoma in the primary care population. Well differentiated as referent category. The figures show relative risks ratios (RRR) with 95% confidence interval (N = 21,215). Note: association estimates (relative risk ratios) were obtained from multinomial logistic regression models, with well differentiated as referent group and all listed variables in the models and controlled for BMI, geographical regions in England and lung related co-morbidities (tuberculosis, COPD and asthma) among lung adenocarcinoma patients. The referent category of the covariates is indicated as a diamond symbol on the first row of each covariate. Age group refers to age at diagnosis. The X axis is displayed on a logarithmic scale. RRRs for ethnicity in the not recorded category are omitted, and can be found in Supplementary Table S9.

while Chinese and Indian populations have consistently shown lower smoking prevalence. However, within these groups, there have been significant gender differences, with a much higher smoking prevalence in men than women.[22] This gender gap in smoking prevalence has contributed to the observed differences in lung cancer incidence across these ethnicities.

Furthermore, we found important differences in the diagnostic characteristics of lung cancer between ethnic groups. Compared with white men, younger women and those from Indian, Caribbean, Black African, Chinese and other Asian backgrounds, were more likely to be diagnosed with adenocarcinoma than squamous cell carcinoma despite adjustment for smoking status. This is an important new insight in additional to existing evidence.[24,25] We also found that despite adjustment for multiple factors, those diagnosed with lung adenocarcinoma from the most deprived backgrounds were at a greater risk of being diagnosed with moderate or poorly differentiated histological grading than those from the most affluent backgrounds. In addition, men and current cigarette smokers were at a higher relative risk of being diagnosed with more advanced stages of lung cancer compared to women and those who never smoked.

Our findings on differences between ethnic groups for histological cancer types, despite adjustment for smoking, adds novel and strong evidence to the existing body of knowledge suggesting that these disparities might not only be attributed to smoking habits but could also be influenced by factors such as ethnicity/genetic predispositions, differences in nicotine metabolism, and sociodemographic variables.[26] The rise in lung adenocarcinoma has been linked to an increase in low tar cigarette use, with deeper inhalation possibly leading carcinogens to areas prone to adenocarcinomas.[25] These findings, coupled with notable gender and racial disparities—adenocarcinoma being more common in women and varied cancer risk factors between white and non-white smokers—underscores the importance of taking into account the unique ethnic and socioeconomic factors in different ethnic groups.

This information is important for radiologists who report on lung cancer scans and for potentially for clinicians delivering smoking prevention and cessation services, highlighting the need for targeted lung cancer prevention and treatment strategies that take into account these multifaceted influences.

Taken together, our results have implications not only for targeting smoking prevention and cessation interventions in an accessible way, but also ensuring equitable delivery of the new lung cancer screening programme especially for women, those from ethnic minority groups and deprived areas to avoid exacerbating health inequalities. The NHS England Targeted Lung Health Check programme, now the national screening programme, began by targeted geographic areas with high incidence and mortality, reflecting the intention to start in the areas most in need. These areas have high levels of socioeconomic deprivation. Local programmes have materials adapted for ethnic minorities and data are being collected on participation rates.

Our findings on lung cancer incidence rates align closely with those reported by Cancer Research UK (CRUK).[27] However, our analyses extend these by adding more details on ethnicity and sociodemographic factors, covering the years 2005–2019, up to the onset of the COVID pandemic lockdown in the UK. We observed gender-specific trends, noting a slight increase in lung cancer incidence among females and a decrease among males over the last decade. Additionally, our analysis reinforces evidence of a significant socioeconomic disparity; lung cancer rates are over twice as high in the most deprived neighbourhoods compared to the more affluent.

Additionally, our study identified a decline in age-standardised lung cancer incidence rates among both male and female never-smokers, with a more pronounced decrease observed in males. These findings are in line with a recent UK study that reported decreasing lung cancer rates among never-smokers of both sexes over the past two decades.[28] Similarly, data from various global cohorts reveal no significant temporal trends in lung cancer incidence and mortality among never-smokers.[29] Nevertheless, some studies have noted rising lung cancer rates—particularly adenocarcinoma—among never-smokers in certain regions, particularly in East Asia and among women.[29] These observations underscore the need for more focused research on non-smoking-related risk factors for lung cancer.

Our study, based on a large, nationally representative, diverse population substantially improves on previous studies limited by small numbers[16,17] or regional populations in Leicester, London or the South-East of England[14] Another study found no difference between ethnic groups for histology or stage at diagnosis, but this was limited as it based on just 423 highly selected patients recruited from a tertiary referral centre which didn't accept referrals of patients with early-stage disease.[17] The authors concluded there was no need to tailor health advice or interventions by ethnicity although it is likely that their study was substantially under-powered. In contrast, our population-based study, is more than 200 times the size and has sufficient numbers to report robust analyses by ethnic group and has therefore reached a different conclusion between ethnic groups and histology. Furthermore, the majority of previous studies were conducted over 20 years ago and the findings are not generalisable to the ethnically diverse contemporaneous populations.

To the best of our knowledge, this study is the largest to date, spanning over 15 years, set to offer comprehensive insights into the variations in lung cancer incidence and diagnostic factors in the UK, with a focus

on disparities related ethnicity and sociodemographic factors, aiming to quantify and address existing health inequalities in lung cancer. The strengths of our study include large size, recent data, and long duration of follow up, representativeness, minimisation of selection, recall and respondent bias. The use of linked hospital, mortality and cancer registry data has increased ascertainment of diagnoses of lung cancer. The linkage to the cancer registry has enabled the inclusion of more granular information on histological type, stage and grade of diagnosis than previous studies, although we acknowledge that there are still varying levels of missing data on cancer histology subtypes, grades, and stage at diagnosis in the linked data. Furthermore, in contrast to previous localised studies in the UK that overlooked smoking status, our study accounted for its crucial role in lung cancer, considering the varying impacts of smoking across different ethnic groups on lung cancer outcomes. However, we acknowledge that different ethnic groups may have diverse smoking intensities and frequencies, which may not be fully captured by simple categorical adjustments. This limitation could influence some of the observed ethnicity effects, particularly in groups with distinct smoking habits.

Our study faced limitations such as missing data for ethnicity, Townsend quintile, smoking status, BMI, and the outcomes on lung cancer histological types, stage and grade. To mitigate the impact on missingness, we introduced a "not recorded" category for all missing exposure and outcome variables, to ensure all available data were included in the analysis to minimise data loss. This approach allows us to include participants whose data might otherwise be excluded due to missing values, which helps maintain the overall sample size and statistical power of the study. However, we recognise that this could introduce potential biases of estimates, especially when embedding trends over time in the "not recorded" categories in study outcomes that vary over follow-up time, could potentially bias the estimates by affecting the distribution of these cases. We observed a decline in lung cancer incidence within the unrecorded ethnicity category, particularly among women (Supplementary Table S1), likely due to improved data categorisation and shifts in healthcare access among minority populations. However, this trend suggests complex relationships between ethnicity and socioeconomic factors that warrant further investigation. We performed multinomial logistic models on the total lung cancer cohort for robustness and to identify the main effects of key variables. Additional analyses including interaction terms between ethnicity and sex, deprivation, and smoking status revealed wide confidence intervals overall due to small sample sizes in minority ethnic groups and missing data. This limited statistical power to detect potential interaction effects in one of the models. Also, our study design has not enabled us to explore the causes of health inequalities nor suggest potential interventions to address them.

The national roll-out of the UK screening programme is currently one of the best globally in terms of coverage of the eligible population, facilitated by the NHS. It is an exemplar of operation delivery of a targeted screening programme. Addressing disparities in coverage due to avoidable factors is important. There has already been considerable activity in this area, and it will be important to ensure data is collected to ensure minority or higher-risks groups are not disadvantaged and provided with equitable access to services. Internationally it is likely that similar health inequalities may also exist in other countries and even if not, the introduction of a new screening programme is likely to generate inequalities as has been repeatedly shown in other clinical areas and populations in accordance with the inverse equity hypothesis.[6] Hence, we believe our findings are likely to be of interest to clinicians and policy makers internationally.

Contributors

DTHC contributed to the development of the research questions, literature review, data manipulation, data analysis, wrote the first draft of the paper. FVG (Fergus V Gleeson, Nuffield Department of Primary Care Health Sciences and Department of Oncology, University of Oxford) and JH-C secured the funding for this study. FVG is the chief investigator of the DART project, and JH-C is the joint package lead (WP6—primary care, population health, and health economics). WL contributed to the study conceptualisation for the overall WP6, data specification for data extraction, drafted the whole research proposal and original statistical analysis plan, worked on the patient and public involvement (PPI) and ethical and scientific approval for WP6. JH-C initiated this analysis, contributed to the development of the research question, organised governance approvals and the data extraction, contributed to interpretation of the results and drafting of the manuscript. JH and CC contributed to review and editing of the original draft. JH-C and CC contributed to the statistical analysis and interpretation. All authors commented critically on drafts of the manuscript and approved the final submission.

Data sharing statement

To guarantee the confidentiality of personal and health information only the authors have had access to the data during the study in accordance with the relevant licence agreements. Access to the QResearch data is according to the information on the QResearch website (www.qresearch.org).

Declaration of interests

JHC reports grants from National Institute for Health Research, John Fell Oxford University Press Research Fund, Cancer Research U.K. (C5255/A18085), Wellcome Institutional Strategic Support Fund (204826/Z/16/Z) and other research councils, during the conduct of the study. JHC is an unpaid director of QResearch, a not-for-profit organisation which is a partnership between the University of Oxford and EMIS Health who supply the QResearch database used for this work. Until 09 August 2023, JHC had a 50% shareholding in ClinRisk Ltd, co-owning it with her husband, who was an executive director. On 9th August 2023, 100% of the share capital was donated to Endeavour Health Care Charitable Trust and the company renamed to Endeavour Predict Ltd. JHC is an unpaid consultant to Endeavour Predict Ltd and her husband is a non-executive director to cover the transition. The company licences software both to the private sector and to NHS bodies or bodies that provide services to the NHS (through GP electronic health record providers, pharmacies, hospital providers and other NHS

providers). This software implements algorithms (including QRISK3) developed from access to the QResearch database during her time at the University of Nottingham. CC reports receiving personal fees from ClinRisk Ltd, outside this work. DRB reports receiving payments for speaking and advice for Astra Zeneca, Roche and MDS.

Acknowledgements

We acknowledge the contribution of EMIS practices who contribute to QResearch and EMIS Health and the Universities of Nottingham and Oxford for expertise in establishing, developing or supporting the QResearch database. This project involves data derived from patient-level information collected by the NHS, as part of the care and support of cancer patients. The Hospital Episode Statistics data, cancer registry and civil registration mortality data are used by permission from NHS England who retain the copyright in that data. NHS England bears no responsibility for the analysis or interpretation of the data.

Funding: The DART (The Integration and Analysis of Data using Artificial Intelligence to Improve Patient Outcomes with Thoracic Diseases) project is funded by Innovate UK, UK Research and Innovation, grant reference: 40255). QResearch® received funding from the NIHR Biomedical Research Centre (Oxford), John Fell Oxford University Press Research Fund, Cancer Research UK (Grant number C5255/A18085), through the Cancer Research UK Oxford Centre, the Oxford Wellcome Institutional Strategic Support Fund (204826/Z/16/Z).

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.lanepe.2024.101124.

References

1 Leiter A, Veluswamy RR, Wisnivesky JP. The global burden of lung cancer: current status and future trends. *Nat Rev Clin Oncol*. 2023;20(9):624–639.
2 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(3):209–249.
3 National Lung Screening Trial Research Team, Aberle DR, Adams AM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*. 2011;365(5):395–409.
4 de Koning HJ, van der Aalst CM, de Jong PA, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. *N Engl J Med*. 2020;382(6):503–513.
5 NHS UK. Adult screening programme: lung cancer 2022 [cited 2023 August 30]. Available from: https://view-health-screening-recommendations.service.gov.uk/lung-cancer/.
6 Victora CG, Joseph G, Silva ICM, et al. The inverse equity hypothesis: analyses of institutional deliveries in 286 national surveys. *Am J Public Health*. 2018;108(4):464–471.
7 Singh GK, Jemal A. Socioeconomic and racial/ethnic disparities in cancer mortality, incidence, and survival in the United States, 1950-2014: over six decades of changing patterns and widening inequalities. *J Environ Public Health*. 2017;2017:2819372.
8 Dwyer LL, Vadagam P, Vanderpoel J, Cohen C, Lewing B, Tkacz J. Disparities in lung cancer: a targeted literature review examining lung cancer screening, diagnosis, treatment, and survival outcomes in the United States. *J Racial Ethn Health Disparities*. 2023;11(3):1489–1500.
9 Houston KA, Mitchell KA, King J, White A, Ryan BM. Histologic lung cancer incidence rates and trends vary by race/ethnicity and residential county. *J Thorac Oncol*. 2018;13(4):497–509.
10 Kish JK, Yu M, Percy-Laurry A, Altekruse SF. Racial and ethnic disparities in cancer survival by neighborhood socioeconomic status in Surveillance, Epidemiology, and End Results (SEER) Registries. *J Natl Cancer Inst Monogr*. 2014;2014(49):236–243.
11 Zavala VA, Bracci PM, Carethers JM, et al. Cancer health disparities in racial/ethnic minorities in the United States. *Br J Cancer*. 2021;124(2):315–332.
12 Moss JL, Pinto CN, Srinivasan S, Cronin KA, Croyle RT. Persistent poverty and cancer mortality rates: an analysis of county-level poverty designations. *Cancer Epidemiol Biomarkers Prev*. 2020;29(10):1949–1954.
13 Gadgeel SM, Severson RK, Kau Y, Graff J, Weiss LK, Kalemkerian GP. Impact of race in lung cancer: analysis of temporal trends from a surveillance, epidemiology, and end results database. *Chest*. 2001;120(1):55–63.
14 Jack RH, Davies EA, Møller H. Lung cancer incidence and survival in different ethnic groups in South East England. *Br J Cancer*. 2011;105(7):1049–1053.
15 Bennett VA, Davies EA, Jack RH, Mak V, Møller H. Histological subtype of lung cancer in relation to socio-economic deprivation in South East England. *BMC Cancer*. 2008;8(1):139.
16 Nimako K, Gunapala R, Popat S, O'Brien ME. Patient factors, health care factors and survival from lung cancer according to ethnic group in the south of London, UK. *Eur J Cancer Care*. 2013;22(1):79–87.
17 Smith LK, Peake MD, Botha JL. Recent changes in lung cancer incidence for south Asians: a population based register study. *BMJ*. 2003;326(7380):81–82.
18 Cranford HM, Koru-Sengul T, Lopes G, Pinheiro PS. Lung cancer incidence by detailed race-ethnicity. *Cancers*. 2023;15(7):2164.
19 Nicholson AG, Tsao MS, Beasley MB, et al. The 2021 WHO classification of lung tumors: impact of advances since 2015. *J Thorac Oncol*. 2022;17(3):362–387.
20 Townsend P. Deprivation. *J Soc Pol*. 1987;16:125–146.
21 Pace M, Lanzieri G, Glickman M, et al. *Revision of the European standard population*. Report of the Eurostat's task force (Publications Office of the European Union, 2013); 2013:20.
22 Action on Smoking and Health -ASH. Tobacco and ethnic minorities. cited 2024. Available from: https://ash.org.uk/uploads/Tobacco-and-Ethnic-Minorities-Fact-Sheet-v3.pdf?v=1712223258; 2024.
23 Chen DT-H, Girvalaki C, Mechili EA, Millett C, Filippidis FT. Global patterns and prevalence of dual and poly-tobacco use: a systematic review. *Nicotine Tob Res*. 2021;23(11):1816–1820.
24 Delon C, Brown KF, Payne NWS, Kotrotsios Y, Vernon S, Shelton J. Differences in cancer incidence by broad ethnic group in England, 2013–2017. *Br J Cancer*. 2022;126(12):1765–1773.
25 Wynder EL, Muscat JE. The changing epidemiology of smoking and lung cancer histology. *Environ Health Perspect*. 1995;103(suppl 8):143–148.
26 Anand S, Vikramdeo KS, Singh S, Singh AP, Dasgupta S. Racial disparities in the genetic landscape of lung cancer. *Cancer Health Disparities*. 2022;6:210.
27 Cancer Research UK. Bowel cancer statistics 2022 [cited 2024 20 May]. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer#heading-Zero.
28 Rait G, Horsfall L. Twenty-year sociodemographic trends in lung cancer in non-smokers: a UK-based cohort study of 3.7 million people. *Cancer Epidemiology*. 2020;67:101771.
29 Thun MJ, Hannan LM, Adams-Campbell LL, et al. Lung cancer occurrence in never-smokers: an analysis of 13 cohorts and 22 cancer registry studies. *PLoS Med*. 2008;5(9):e185.