Environmental correlates of chronic obstructive pulmonary disease in 96 779 participants from the UK Biobank: a cross-sectional, observational study

Chinmoy Sarkar, Bing Zhang, Michael Ni, Sarika Kumari, Sarah Bauermeister, John Gallacher, Chris Webster

Lancet Planet Health 2019; 3: e478–90

Healthy High Density Cities Lab, HKUrbanLab (C Sarkar PhD, B Zhang MD, M Ni MD, S Kumari MPhil, Prof C Webster DSc) **and School of Public Health** (C Sarkar, M Ni), **The University of Hong Kong, Pokfulam, Hong Kong Special Administrative Region, China; and Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford, UK** (S Bauermeister PhD, Prof J Gallacher PhD)

Correspondence to:
Dr Chinmoy Sarkar, Healthy High Density Cities Lab, HKUrbanLab, The University of Hong Kong, Pokfulam, Hong Kong Special Administrative Region, China
csarkar@hku.hk

Summary

Background The role of environmental exposures in chronic obstructive pulmonary disease (COPD) remains inconclusive. We examined the association between environmental exposures ($PM_{2·5}$, greenness, and urbanicity) and COPD prevalence using the UK Biobank cohort data to identify key built environment correlates of COPD.

Methods In this cross-sectional, observational study we used baseline data for UK Biobank participants. Included participants were aged 39 years and older, white, had available spirometry data, and had complete data for phenotypes and exposures. COPD was defined by spirometry with the 2017 Global Initiative for Chronic Obstructive Lung Disease criteria. Environmental exposures were $PM_{2·5}$ derived from monitoring data and interpolated using land-use regression at the participants' geocoded residential addresses. Built environment metrics of residential greenness were modelled in terms of normalised difference vegetation index from remotely sensed colour infrared data within a 500 m residential catchment, and an urbanicity index derived from spatial analyses and measured with a 1 km buffer around each participant's residential address. Logistic regression models examined the associations between environmental exposures and COPD prevalence adjusting for a range of confounders. Subgroup analyses by urbanicity and effect modification by white blood cell count as an inflammatory marker were also done.

Findings We assessed 96 779 participants recruited between April 4, 2006, and Oct 1, 2010, of which 5391 participants had COPD with a prevalence of 5·6%. Each 10 μg/m³ increment in ambient $PM_{2·5}$ exposure at a participant's residential location was associated with higher odds of COPD (odds ratio 1·55, 95% CI 1·14–2·10). Among the built environment metrics, urbanicity was associated with higher odds of COPD (1·05, 1·01–1·08 per interquartile increment), whereas residential greenness was protective, being associated with lower odds of COPD (0·89, 0·84–0·93 for each interquartile increment in greenness). The results remained consistent in models of COPD defined as per lower limit of normal criteria. The highest quartile of white blood cell count was associated with lower lung function and higher COPD risk with a significant interaction between $PM_{2·5}$ and white blood cell count only in the model of lung function (p=0·0003).

Interpretation In this study of the built environment and COPD, to our knowledge the largest done in the UK, we found that exposure to ambient $PM_{2·5}$ and urbanicity were associated with a higher risk of COPD. Residing in greener areas, as measured by normalised difference vegetation index, was associated with lower odds of COPD, suggesting the potential value of urban planning and design in minimising or offsetting environmental risks for the prevention and management of COPD.

Funding University of Hong Kong, UK Biobank, and UK Economic & Social Research Council.

Introduction

Chronic obstructive pulmonary disease (COPD) is a global public health challenge affecting approximately 174·5 million people, accounting for 3·2 million deaths in 2015.[1] It has emerged as the tenth leading cause of years of life lost and eighth in terms of years lived with disability globally.[2,3] A 2016 UK study estimated that COPD prevalence will increase from 1·79% to 2·19% in England and 2·03% to 2·20% in Scotland from 2011–30, with a projected total health cost of £2·53 billion by 2030.[4]

Exposure to air pollution is one of the main triggers for symptoms of COPD and its exacerbation.[5–9] Outdoor $PM_{2·5}$, in particular, poses considerable health risks, being the fifth leading risk factor for death globally.[10] $PM_{2·5}$ can reach the lung's gas exchange zone, accumulate in the alveolar regions, and penetrate through the respiratory barrier. The soluble components can then enter the bloodstream in the form of ultrafine particulate matter via the alveolar capillaries and the insoluble components can be deposited as sediments in lungs. COPD development and its exacerbations have been attributed to airflow obstruction, inflammation, oxidative stress, immune dysfunction, altered airway epithelial structure, and the microbiome.[11,12]

Several studies have examined the long-term effects of exposure to $PM_{2·5}$ and COPD and related risks. WHO's

Research in context

Evidence before this study

We searched PubMed, MEDLINE, Scopus, and Google or Google Scholar for research papers published in English from database inception until July 2, 2019, with search terms "particulate matter", "$PM_{2·5}$" OR "green space" OR "urbanicity"; AND "COPD" AND "chronic obstructive pulmonary disease". Additionally, we manually searched the reference lists of related papers. Several studies have established positive associations between exposure to $PM_{2·5}$ and risks of COPD prevalence, its exacerbation, and COPD-related mortality. There has been very little evidence linking COPD with greenness and urbanicity. The pathway from exposure to COPD development has been hypothesised to proceed via several mechanisms including inflammation, oxidative stress, immune dysfunction, altered airway epithelial structure, and the microbiome within the lung. Yet, the evidence regarding $PM_{2·5}$ exposure and COPD prevalence has been mostly suggestive and is far from conclusive. Most studies have examined $PM_{2·5}$ exposure in isolation without accounting for other potentially important built environment attributes. Many of the studies have used self-reported doctor-diagnosed COPD, whereas others used hospital visits for COPD or health insurance records of mortality due to COPD, which are methods that have reduced objectivity in capturing prevalent COPD cases. Large-scale evidence has thus far been scarce with most studies done on a homogeneous population with limited particulate exposure variability.

Added value of this study

To our knowledge, the present study is the first of its kind to use a UK-wide dataset of unprecedented size and heterogeneity (n=96 779) to examine the associations between COPD prevalence and multiple environmental exposures of $PM_{2·5}$, residential greenness, and urbanicity after adjusting for a range of covariates and confounders. $PM_{2·5}$ was modelled at location of geocoded residence. Residential greenness was modelled from 0·50 m resolution colour infrared data (enabling higher precision than has previously been possible), while urbanicity has been objectively measured by spatial analyses, comprising indices of density and street movement density. Both exposures were measured around street catchments of participants' geocoded residential addresses. We used rigorous criteria for assessment of COPD cases using spirometry and tested models with an alternative definition of COPD as a sensitivity analysis. Additionally, we developed dose-response curves showing variation in lung function and odds of COPD across $PM_{2·5}$ and green exposure continuums. Additional sensitivity tests comprised examining effects among non-smokers and subgroups of urbanicity quartiles. Our analyses are the first to our knowledge to adjust for haematological inflammatory markers and to test for effect modification by white blood cell count (as an inflammatory marker).

Implications of all the available evidence

Overall, our study reported a 54·8% higher odds of COPD per 10 μg/m³ increment in $PM_{2·5}$ exposure and 4·6% higher odds per IQR increase in urbanicity. The effects were protective for residential greenness with 11·4% lower odds of COPD per IQR increment in normalised difference vegetation index greenness. The potential value of urban planning and design such as appropriate green infrastructure and urban ventilation corridors in minimising or offsetting environmental risks in respect to the prevention and management of COPD needs to be further explored.

SAGE cohort reported positive associations between odds of self-reported doctor diagnosed COPD and $PM_{2·5}$ exposure.[13] Studies from China from 2018 and 2017 used spirometry to diagnose COPD and reported positive associations between exposure to $PM_{2·5}$ and COPD prevalence.[14,15] On the other hand, a Europe-wide study reported null associations between $PM_{2·5}$ exposure and both prevalent and incident COPD outcomes.[16] Several US studies used Medicare beneficiary data, which reported a positive association between long-term $PM_{2·5}$ exposure within an individual's ZIP code of residence and risk of COPD mortality.[17,18] Relatedly, evidence exists for the detrimental effects of $PM_{2·5}$ on lung function[19,20] and COPD exacerbations.[21] Chronic bronchitis, a phenotype of COPD, has also been reported to be positively associated with $PM_{2·5}$ in a cohort of US women,[22] whereas two studies found null associations.[23,24] The links between short-term exposure to $PM_{2·5}$ and COPD-related emergency department visits, hospital admissions, and mortality have also been supported in time-series analyses.[25,26]

The influence of built environment factors in respiratory health has been acknowledged.[27] Few studies indicate protective effects of residential greenness on respiratory health,[28] and COPD[29] in particular. Evidence for asthma, allergic rhinitis, and aeroallergen sensitisation (thus far mostly among children) has been inconsistent[30] with some studies reporting protective effects,[31] as well as null,[32] or non-beneficial associations.[33,34]

The evidence associating environmental exposures with COPD thus far has been inconclusive. Most studies have only examined $PM_{2·5}$ exposures without accounting for other potentially important built environment attributes such as residential greenness and urbanicity. Several studies have used self-reported doctor-diagnosed COPD, whereas others used hospital visits for COPD or health insurance records of mortality due to COPD, with very few studies using objectively defined COPD by spirometry, meaning a high likelihood of underestimation of the true burden. Many of the studies have been of small scale, insufficiently powered, and done on a homogeneous population with a narrow range of exposures.

With the aim of understanding environmental correlates of COPD prevalence, we used UK Biobank health datasets to examine the odds of COPD prevalence

with respect to environmental correlates of $PM_{2·5}$, greenness, and urbanicity after adjusting for individual-level covariates and confounders. We also examined effects among non-smokers and across categories of urbanicity with white blood cell count used as an inflammatory marker.

Methods

Study design and participants

We did a cross-sectional, observational study using data from the UK Biobank, the largest European Biobank developed to study the environmental, social, and genetic causes of chronic diseases.[35] The UK Biobank assessed 502 682 participants recruited from National Health Service patient registers and residing within approximately 25 miles of assessment centres located in 22 cities of England, Wales, and Scotland over the baseline period April 4, 2006, and Oct 10, 2010. More than 86·2% of the cohort participants resided in urban areas. Extensive data were collected through a set of questionnaires on sociodemographic, lifestyle, and psychosocial factors, and medical history. Data collection comprised touchscreen questionnaires involving direct participant responses into the data entry system at one of the assessment centres and verbal interviews done by a trained member of staff. Phenotypic characterisations involved anthropometry, biosampling (blood, urine, and saliva), spirometry, imaging, and cognitive function assessments.

The present study was restricted to participants of white European ancestry with available spirometry data (forced expiratory volume in 1 s [FEV_1]; forced vital capacity [FVC] measurements meeting the American Thoracic Society or European Respiratory Society criteria), and complete data on phenotypes and exposures. The study's analytical sample comprised 96 779 participants across 21 assessment centres. The UK Biobank received ethical approval from the National Health Service National Research Ethics Service (Ref:11/NW/0382) and all study procedures adhered to the World Medical Association Declaration of Helsinki ethical principles for medical research. We submitted a research proposal and data access application that was approved by the UK Biobank Scientific Committee (application number 26492). Participants provided electronically signed informed consent. The detailed study protocol including scientific rationale, sampling, and participant selection criteria has been described previously.[36]

Procedures

Participants completed breath spirometry measurements using a Vitalograph Pneumotrac 6800 recording between two to three blows (each lasting for at least 6 s) within a period of about 6 min. The reproducibility of the first two blows was compared by the Spirometer software and if the difference between FVC and FEV_1 was 5% or less, it was deemed acceptable and a third blow was not required. Participants with previous contraindications including chest infection (eg, influenza, bronchitis, severe cold, pneumonia) in the past month; history of detached retina, heart attack, or surgery to eyes, chest, or abdomen in the past 3 months; history of a collapsed lung, pregnancy (first or third trimester), and those currently on medication for tuberculosis were excluded.[37] Post-bronchodilator was not available for the study and participants were not instructed to withhold any drug treatment before spirometry. COPD was defined as per stage II plus classification of airflow limitation, as per the Global Initiative for Chronic Obstructive Lung Disease guidelines.[38] Participants fulfilling the spirometric criteria of FEV_1/FVC ratio less than 0·7 and percentage predicted FEV_1 less than 80% were defined as COPD cases. Predicted percentage FEV_1 was calculated as per UK Biobank's protocol. COPD cases with simultaneous doctor-diagnosed and self-reported asthma were excluded from the analyses.

The air pollution measures in the UK Biobank were provided by the Small Area Health Statistics Unit and were developed as a part of the BioSHaRE-EU Environmental Determinants of Health Project. Residential exposure to ambient $PM_{2·5}$ was obtained from UK Biobank's linked air pollution exposure data obtained from the European Study of Cohorts for Air Pollution Effects (ESCAPE) project. Ambient $PM_{2·5}$ was measured between Jan 26, 2010, and Jan 18, 2011, across 20 sites each in the two study areas of Thames Valley London and Oxford and smaller towns using Harvard impactors. To consider the diverse factors responsible for pollution variability, one site was located outside the urban area and not influenced by traffic-related emissions, 12 sites were in urban areas but at least 50 m away from a major road, while seven were street sites located at building facades of residences adjacent to streets with traffic intensities of around 10 000 vehicles per day or more. Measurements were taken three times annually for 14 days in the cold, warm, and intermediate seasons of the year. To account for temporal variability, one centrally located site was deployed as the reference site where measurements were taken over the entire year. Mean annual $PM_{2·5}$ concentrations were generated for each site after correcting for temporal variability. The ESCAPE project was able to achieve a correlation (r^2) between the measured annual mean $PM_{2·5}$ concentrations and other pollutant components for the London and Oxford areas; 0·84 (nitrogen dioxide), 0·86 (PM_{10}), and 0·79 ($PM_{2·5}$ absorbance). Individual-level $PM_{2·5}$ exposures were generated at each geocoded participant address with land-use regression models using land use, street network, and road traffic variables.[39] The individual-level estimates have been validated to be reliable up to 400 km away from the Greater London area, and so all UK Biobank participant addresses outside this limit were assigned missing values.[40] To consider the effects of reliable accumulated exposures, we restricted all analyses

For **UK Biobank's protocol** see http://biobank.ndph.ox.ac.uk/crystal/field.cgi?id=20153

For more on the **Small Area Health Statistics Unit** see http://www.sahsu.org/

For **BioSHaRE-EU** see http://www.bioshare.eu/

For the **ESCAPE project** see http://www.escapeproject.eu/

to participants who resided in their current address for 3 or more years.

Residential greenness and urbanicity, derived from the UK Biobank Urban Morphometric Platform (UKBUMP), were used as built environment exposures. UKBUMP was developed by some of the authors of this Article (CS, SK, JG, CW) based at the University of Hong Kong (Hong Kong) and University of Oxford (Oxford, UK). UKBUMP is a spatial database of objectively modelled built environment metrics of density, greenness, destination proximity, street-level accessibility, and physical environment around residential activity neighbourhoods of participants and developed to identify environmental predictors of chronic disease.[41]

Residential greenness was measured in terms of normalised difference vegetation index (NDVI), an objective index of relative overall green vegetation or biomass derived from the spectral reflectance values of image pixels in remotely sensed data. Chlorophyll in healthy vegetation absorbs radiation in the visible red region (630–690 nm) of the electromagnetic spectrum and reflects radiation in the near-infrared region (760–900 nm) and this differential in absorbance and reflectance wavelengths is used as a proxy for green quality and intensity. Index scores range between –1 to +1 with higher values indicating dense green vegetation. NDVI was modelled from a series of very high resolution (0·50 m by 0·50 m) Bluesky Colour Infrared imagery derived from specially developed sensors mounted underneath a survey aircraft (appendix pp 3–5). As per established protocols,[42] data preparation involved mosaicking summer-time image tiles of the study areas around UK Biobank assessment centres collected over similar temporal scales (across the baseline phase of the UK Biobank study) to avoid potential temporal mismatch and effect of seasonal variability in greenness. We were able to exclude all large water bodies before the analyses of NDVI. We measured residential greenness as mean NDVI values within a 500-m catchment radius of geocoded UK Biobank participants' dwellings as per previous studies.[43,44] The NDVI greenness was measured for all the participants of 17 (77%) of 22 assessment centres (70% of the total number of individuals) based on colour infrared imagery data available for the study.

See Online for appendix

The land use and street-level physical accessibility metrics were developed from the Ordnance Survey GB spatial dataset. The density metrics were derived from the UK-wide AddressBase Premium dataset of Ordnance Survey consisting of approximately 36 million valid address point features with 550 different land-use classifications. The spatial analytics involved delineating street catchments around geocoded participant dwellings and measuring densities of more than 200 health promoting or inhibiting land-use destinations. We developed a composite index of urbanicity from the UKBUMP built environment database, which had been tested previously.[43] It was derived from four key indicators: residential density, retail density, walkability (street movement density), and density of public transport measured within 1 km of a geocoded participant dwelling and expressed as:

$$Urbanicity = Z\ score_{resid} + Z\ score_{retail} + Z\ score_{PT} + Z\ score_{walkability}$$

where resid represents the density of residential housing units, retail represents the density of retail units per square km street catchment, and PT represents the density of public transport in units per square km street catchment, while walkability is expressed in terms of street movement potential. The former three metrics are expressed as number of units per km² within the 1 km buffer. Street movement potential was obtained by network modelling of the UK-wide street centreline data for the study area, derived from the Integrated Transport Network layer of the Ordnance Survey database comprising approximately 4 million street segments. The network data was transcribed into an access graph model and the street-level movement potential was modelled in the spatial Design Network Analysis network analysis algorithm.[45] Movement potential is expressed as the simulated counts of movement passing through each link in the network, given its relative position and topological connectivity with other segments within the network. The measure also acts as a proxy for relative accessibility and centrality of a place.

Blood samples were drawn from each participant at baseline with standard procedures. Haematology analysis was done in 4 mL EDTA (edetic acid) vacutainers with four Beckman Coulter LH750 instruments. Haematology analysis comprised measuring a series of blood parameters including full red blood cell and white blood cell counts, proportion and counts of individual white blood cell populations, and proportion of reticulocytes of the red blood cell population. White blood cell count, neutrophil-to-lymphocyte ratio, and eosinophil-to-basophil ratio were used as inflammatory biomarkers, given their established role in the pathogenesis of COPD.[46–48]

Individual-level sociodemographic covariates comprised age, sex, highest qualification, and employment status. Highest qualification was a five-factor variable (coded as none; O levels, General Certificate of Secondary Education, or Certificate of Secondary Education; A levels or AS levels; National Vocational Qualification, Higher National Diploma, Higher National Certificate, or other professional qualification; college or university degree), while employment status was a three-factor variable (coded as employed; retired; and unemployed, home maker, others). Average total income before tax, measured at household-level was a four-level factor (<£18 000, £18 000–30 999, £31 000–51 999, ≥£52 000). Among the individual lifestyle-level risk factors, smoking was expressed as a five-factor variable (never-smokers, current or past occasional smokers, pack years <10,

10–19, and ≥20). Alcohol intake frequency was coded as a four-factor variable (never or occasional, 1–2 times per week, 3–4 times per week, daily or almost daily). Physical activity was expressed as log-transformed metabolic equivalent of task h per week from the self-reported International Physical Activity Questionnaire short form comprising weekly walking, moderate physical activity, and vigorous physical activity components. Residential tenureship was a household-level three-factor variable (own outright, own with mortgage, rented). Townsend deprivation index scores (coded in quintiles) were used as an indicator of area-level socioeconomic status with a higher score indicative of lower neighbourhood socioeconomic status. Biological risk factors comprised anthropometrics and morbidity-related variables assessed for each participant and included standing height, body-mass index (BMI) status (coded as three-factor variable as BMI <25; BMI ≥25 to <30; BMI ≥30 kg/m²), parent's COPD status (no COPD, one parent with COPD, both parents with COPD), diabetes status (no diabetes, prevalent diabetes) and cardiovascular disease status (none; high blood pressure; or heart attack, angina, or stroke; both high blood pressure and heart attack, angina, or stroke).

Statistical analysis

We used logistic regression models with robust variance estimates. Separate models were developed to examine the associations between the environmental risk factors of $PM_{2.5}$, urbanicity, residential greenness, and COPD after adjusting for a range of covariates and confounders (selection informed a priori by literature[49]) to include pertinent sociodemographics, lifestyle variables, neighbourhood socioeconomic status, anthropometrics, comorbidities, and haematological biomarkers. Models 1, 2, and 3 represent multiple adjusted models developed for $PM_{2.5}$, urbanicity, and residential greenness, respectively. The initial model building exercise consisted of assessing correlations between the exposure variables, doing bivariate analyses, and subsequently sequentially introducing blocks of risk factors to examine their confounding effects on point estimates, significance, collinearity (assessed with variance inflationary factor), and fit statistic to ensure a parsimonious fit. The final models selected comprised the selected blocks of sociodemographic, lifestyle, neighbourhood-level, biological, and haematological risk factors.

As sensitivity analysis, we tested our results using an alternative definition of COPD based on lower limit of normal criteria. Participants were assigned a COPD case if they fulfilled the spirometric criteria of FEV_1/FVC ratio less than the lower limit of normal, which is less than the lower fifth percentile tail of the normal distribution of mean predicted FEV_1/FVC in a reference healthy population. We used the Hankinson's equation to estimate the predicted values of FEV_1/FVC.[50] We further tested for non-linearity in the associations and developed dose-response curves showing variation in lung function (defined as the ratio of FEV_1/FVC), and odds of COPD across the continuum of $PM_{2.5}$ and green exposures, using restricted cubic spline models with Harrel's knots. To further examine potential pathways from $PM_{2.5}$ exposure to COPD through an inflammatory haematological biomarker, we introduced an interaction term between

	Individuals without COPD (n=91 388)	Individuals with COPD (n=5391)	Analytical sample (N=96 779)	OR (95% CI)*
Age (years)	55·7 (7·8)	60·0 (6·8)	56·2 (7·8)	1·08 (1·07–1·08)
Forced expiratory volume in 1 s, L	3·0 (0·7)	1·9 (0·5)	2·9 (0·7)	NA
Forced vital capacity, L	3·8 (0·9)	3·1 (0·8)	3·8 (0·9)	NA
Standing height, cm	168·4 (9·0)	169·8 (9·2)	168·4 (9·0)	1·02 (1·01–1·02)
Body-mass index, kg/m²	27·1 (4·5)	27·3 (4·8)	27·1 (4·5)	1·01 (1·00–1·02)
Physical activity (metabolic equivalent of task h per week)	46·8 (49·0)	47·8 (53·7)	46·9 (49·2)	0·95 (0·92–0·98)
Sex				
Female	53 707 (58·8%)	2417 (44·8%)	56 124 (58·0%)	NA
Male	37 681 (41·2%)	2974 (55·2%)	40 655 (42·0%)	1·75 (1·66–1·85)
Annual household income				
<£18 000	15 544 (17·0%)	1817 (33·7%)	17 361 (17·9%)	NA
£18 000–30 999	22 159 (24·2%)	1507 (28·0%)	23 666 (24·5%)	0·58 (0·54–0·62)
£31 000–51 999	25 655 (28·1%)	1147 (21·3%)	26 802 (27·7%)	0·38 (0·35–0·41)
≥£52 000	28 030 (30·7%)	920 (17·1%)	28 950 (29·9%)	0·28 (0·26–0·3)
Highest qualification				
None	10 032 (11·0%)	1285 (23·8%)	11 317 (11·7%)	NA
O levels, GCSEs, or CSEs	15 694 (17·2%)	929 (17·2%)	16 623 (17·2%)	0·46 (0·42–0·50)
A levels or AS levels	5488 (6·0%)	272 (5·0%)	5760 (6·0%)	0·39 (0·34–0·44)
NVQ, HND, HNC, or other professional qualification	26 504 (29·0%)	1561 (29·0%)	28 065 (29·0%)	0·46 (0·43–0·50)
College or university degree	33 670 (36·8%)	1344 (24·9%)	35 014 (36·2%)	0·31 (0·29–0·34)
Employment status				
Employed	58 022 (63·5%)	2438 (45·2%)	60 460 (62·5%)	NA
Retired	27 757 (30·4%)	2473 (45·9%)	30 230 (31·2%)	2·12 (2·00–2·25)
Unemployed, home maker, others	5609 (6·1%)	480 (8·9%)	6089 (6·3%)	2·04 (1·84–2·25)
Alcohol intake frequency				
Never or occasional	23 734 (26·0%)	1571 (29·1%)	25 305 (26·2%)	NA
1–2 times per week	23 973 (26·2%)	1247 (23·1%)	25 220 (26·1%)	0·79 (0·73–0·85)
3–4 times per week	23 647 (25·9%)	1137 (21·1%)	24 784 (25·6%)	0·73 (0·67–0·79)
Daily or almost daily	20 034 (21·9%)	1436 (26·6%)	21 470 (22·2%)	1·08 (1·01–1·17)
Smoking status				
Never smokers	52 150 (57·1%)	1530 (28·4%)	53 680 (55·5%)	NA
Current or past occasional smokers	13 161 (14·4%)	511 (9·5%)	13 672 (14·1%)	1·32 (1·20–1·47)
Pack years <10	8135 (8·9%)	309 (5·7%)	8444 (8·7%)	1·30 (1·14–1·47)
Pack years 10–19	7984 (8·7%)	552 (10·2%)	8536 (8·8%)	2·36 (2·13–2·6)
Pack years ≥20	9958 (10·9%)	2489 (46·2%)	12 447 (12·9%)	8·52 (7·97–9·11)
Residential tenureship				
Own outright	49 029 (53·6%)	3234 (60·0%)	52 263 (54·0%)	NA
Mortgage	37 655 (41·2%)	1445 (26·8%)	39 100 (40·4%)	0·58 (0·55–0·62)
Rent	4704 (5·1%)	712 (13·2%)	5416 (5·6%)	2·30 (2·10–2·50)

(Table 1 continues on next page)

	Individuals without COPD (n=91 388)	Individuals with COPD (n=5391)	Analytical sample (N=96 779)	OR (95% CI)*
(Continued from previous page)				
Townsend deprivation				
Quintile 1, low	42 548 (46·6%)	1984 (36·8%)	44 532 (46·0%)	NA
Quintile 2	13 220 (14·5%)	758 (14·1%)	13 978 (14·4%)	1·23 (1·13–1·34)
Quintile 3	12 430 (13·6%)	761 (14·1%)	13 191 (13·6%)	1·31 (1·20–1·43)
Quintile 4	12 936 (14·2%)	873 (16·2%)	13 809 (14·3%)	1·45 (1·33–1·57)
Quintile 5, high	10 254 (11·2%)	1015 (18·8%)	11 269 (11·6%)	2·12 (1·96–2·30)
Cardiovascular problems				
None	68 334 (74·8%)	3282 (60·9%)	71 616 (74·0%)	NA
High blood pressure	19 655 (21·5%)	1541 (28·6%)	21 196 (21·9%)	1·63 (1·53–1·74)
Heart attack, angina, stroke	1689 (1·8%)	245 (4·5%)	1934 (2·0%)	3·02 (2·63–3·47)
High blood pressure and heart attack, angina, stroke	1710 (1·9%)	323 (6·0%)	2033 (2·1%)	3·93 (3·47–4·45)
Diabetes status				
No diabetes	88 062 (96·4%)	5013 (93·0%)	93 075 (96·2%)	NA
Diabetes	3326 (3·6%)	378 (7·0%)	3704 (3·8%)	2·00 (1·79–2·23)
Parent's COPD status				
No COPD	78 543 (85·9%)	4112 (76·3%)	82 655 (85·4%)	NA
One parent with COPD	12 199 (13·3%)	1179 (21·9%)	13 378 (13·8%)	1·85 (1·73–1·97)
Both parents with COPD	646 (0·7%)	100 (1·9%)	746 (0·8%)	2·96 (2·39–3·66)
White blood cell counts, 10^9 cells per L*				
Quartile 1, low (<5·59)	21 535 (23·6%)	732 (13·6%)	22 267 (23·0%)	NA
Quartile 2 (5·59–6·54)	22 610 (24·7%)	984 (18·3%)	23 594 (24·4%)	1·28 (1·16–1·41)
Quartile 3 (6·55–7·70)	23 660 (25·9%)	1337 (24·8%)	24 997 (25·8%)	1·66 (1·52–1·82)
Quartile 4, high (>7·70)	23 583 (25·8%)	2338 (43·4%)	25 921 (26·8%)	2·92 (2·68–3·18)
Neutrophil-to-lymphocyte ratio				
Quartile 1, low (<1·66)	22 822 (25·0%)	1142 (21·2%)	23 964 (24·8%)	NA
Quartile 2 (1·66–2·11†)	23 098 (25·3%)	1195 (22·2%)	24 293 (25·1%)	1·03 (0·95–1·12)
Quartile 3 (2·11‡–2·72)	22 931 (25·1%)	1367 (25·4%)	24 298 (25·1%)	1·19 (1·10–1·29)
Quartile 4, high (>2·72)	22 537 (24·7%)	1687 (31·3%)	24 224 (25·0%)	1·50 (1·38–1·62)
Eosinophil-to-basophil ratio				
Quartile 1, low (<2·17)	22 909 (25·1%)	1299 (24·1%)	24 208 (25·0%)	NA
Quartile 2 (2·17–4·00)	23 495 (25·7%)	1396 (25·9%)	24 891 (25·7%)	1·05 (0·97–1·13)
Quartile 3 (4·01–7·50)	23 267 (25·5%)	1331 (24·7%)	24 598 (25·4%)	1·01 (0·93–1·09)
Quartile 4, high (>7·50)	21 717 (23·8%)	1365 (25·3%)	23 082 (23·9%)	1·11 (1·03–1·20)

Data are mean (SD) or n (%) unless specified. Physical activity (metabolic equivalent of task h per week) was available for 80 395 participants. COPD=chronic obstructive pulmonary disorder. OR=odds ratio. NA=not applicable. GCSE=General Certificate of Secondary Education. CSE=Certificate of Secondary Education. NVQ=National Vocational Qualification. HND=Higher National Diploma. HNC=Higher National Certificate. *Bivariate models of association between COPD defined as per Global Initiative for Chronic Obstructive Lung Disease stage II plus classification and each participant's characteristics. †2·112299. ‡2·11236.

Table 1: **Baseline characteristics**

$PM_{2·5}$ and quartiles of white blood cell count in our restricted cubic spline model examining association of $PM_{2·5}$ exposure with lung function and COPD.

To study potential confounding effects of greenness on the association between COPD and $PM_{2·5}$, we introduced NDVI greenness in our primary analysis. We restricted analyses to non-smokers to understand the potential effects of $PM_{2·5}$, urbanicity, and greenness among non-smokers. We further ran subgroup analyses by quartiles of urbanicity scores to understand the potential impact of urbanicity on the effect estimates in our models of association of COPD with $PM_{2·5}$ and NDVI greenness.

We report odds ratios (ORs) and two-tailed 95% CIs estimated with robust variance estimator (Huber-White sandwich estimator) to account for potential clustering within data. All analyses were done in Stata version 15.1.

Role of the funding source

UK Biobank was involved in data collection. The funders of the study had no role in developing the research questions, study design, data modelling and analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

The analytical sample comprised 96 779 participants recruited between April 4, 2006, and Oct 1, 2010. Valid spirometry data as per American Thoracic Society or European Respiratory Society guidelines were available for 275 897 participants. After restricting analyses to white European ancestry and excluding those with incomplete phenotypes and those not meeting the American Thoracic Society or European Respiratory Society criteria, as well as missing data for pack years of smoking, 208 722 participants remained with valid data for prevalent COPD. After the further exclusion of participants who resided in their current residential address for less than 3 years (n=16 996), and had missing data for $PM_{2·5}$ (n=15 933), sociodemographic and lifestyle level variables (n=27 777), comorbidities (n=15 561), and haematological biomarkers (n=35 676), the study included 96 779 participants. Comparisons of the analytical sample with the full UK Biobank cohort are presented in the appendix (pp 1–2). The sample had high residential stability, with 88·2% of the participants residing in their current address for 5 or more years; 69·3% had for 10 or more years, and mean duration of residence was 18·4 years. The mean age of the participants was 56·2 years (SD 7·8) and 56 124 (58%) of 96 779 were female. The overall spirometry-defined COPD prevalence was 5·6% with 5391 of 96 779 cases and the mean lung function was 0·77 (SD 0·05). Participants reporting more than 20 pack years of cigarette smoking constituted 12·9% (n=12 447) of the analytical sample. The mean $PM_{2·5}$ exposure (μg/m³) was 9·91 (SD 1·03, IQR 1·25). The mean urbanicity index was –0·12 (2·89, 3·13). The mean residential green exposure measured in terms of NDVI metric was 0·18 (0·16, 0·23) and this was available for 77 679 participants. The Pearson's correlation of $PM_{2·5}$ with urbanicity was 0·70 and with residential greenness was –0·06. The correlation of urbanicity with residential greenness was –0·02. The mean inflammatory biomarkers of white blood cell count were $6·9\times10^9$ cells per L (SD $2·2\times10^9$), neutrophil-to-lymphocyte ratio was 2·3 (1·0), and eosinophil-to-basophil ratio was 5·8 (5·9). For descriptive characteristics of the

	Mean (SD)	P_{25}	Median	P_{75}	P_{90}	OR (95% CI)*
$PM_{2·5}$, μg/m³	9·91 (1·03)	9·22	9·87	10·47	11·16	3·95 (3·06–5·10)
Residential density, units per km²	1871·18 (1176·61)	1207·80	1728·13	2247·92	2995·70	1·09 (1·06–1·11)
Retail density, units per km²	43·92 (65·48)	7·23	21·40	54·28	112·67	1·06 (1·04–1·07)
Public transport density, units per km²	21·86 (10·36)	15·08	21·06	27·58	34·44	1·19 (1·16–1·23)
Street-level movement density, ×10⁶	4·54 (5·32)	0·85	2·69	6·35	11·4	1·05 (1·02–1·07)
Urbanicity	–0·12 (2·89)	–1·98	–0·57	1·15	3·25	1·14 (1·11–1·17)
Mean NDVI†	0·18 (0·16)	0·06	0·18	0·29	0·39	0·88 (0·84–0·92)

P_{25} represents the 25th, P_{75} the 75th, and P_{90} the 90th data percentiles. NDVI=normalised difference vegetation index. OR=odds ratio. COPD=chronic obstructive pulmonary disorder. *Bivariate models of association between COPD defined as per Global Initiative for Chronic Obstructive Lung Disease stage II plus classification and environmental risks. †Mean NDVI was available for 77 679 participants with 4325 COPD cases. The estimates represent per 10 μg/m³ increment in $PM_{2·5}$ and one IQR increment in urbanicity and NDVI greenness.

Table 2: **Environmental exposures**

	Model 1 (n_1=96 779)	Model 2 (n_2=94 265)	Model 3 (n_3=77 679)
Environmental exposures			
$PM_{2·5}$, per 10 μg/m³	1·55 (1·14–2·10)	..	..
Urbanicity, 1 km per IQR	..	1·05 (1·01–1·08)	..
Mean NDVI greenness within 500 m, per IQR	..	..	0·89 (0·84–0·93)
Demographics			
Age	1·06 (1·05–1·07)	1·06 (1·05–1·07)	1·06 (1·05–1·07)
Sex			
Female	1 (ref)	1 (ref)	1 (ref)
Male	1·11 (1·02–1·21)	1·11 (1·02–1·22)	1·07 (0·97–1·17)
Household income			
<£18 000	1 (ref)	1 (ref)	1 (ref)
£18 000–30 999	0·85 (0·78–0·92)	0·85 (0·79–0·92)	0·91 (0·83–1·00)
£31 000–51 999	0·75 (0·69–0·82)	0·75 (0·69–0·83)	0·78 (0·70–0·86)
≥£52 000	0·71 (0·64–0·79)	0·71 (0·64–0·79)	0·78 (0·69–0·88)
Highest qualification			
None	1 (ref)	1 (ref)	1 (ref)
O levels, General Certificate of Secondary Education, or Certificate of Secondary Education	0·83 (0·75–0·91)	0·82 (0·75–0·91)	0·77 (0·69–0·86)
A levels or AS levels	0·76 (0·66–0·88)	0·74 (0·63–0·85)	0·70 (0·59–0·82)
National Vocational Qualification, Higher National Diploma, Higher National Certificate, or other professional qualification	0·84 (0·77–0·92)	0·84 (0·77–0·92)	0·80 (0·73–0·88)
College or university degree	0·77 (0·70–0·84)	0·76 (0·69–0·83)	0·73 (0·66–0·81)
Employment status			
Employed	1 (ref)	1 (ref)	1 (ref)
Retired	0·96 (0·89–1·04)	0·95 (0·88–1·03)	0·95 (0·87–1·04)
Unemployed, home maker, others	1·23 (1·10–1·38)	1·22 (1·09–1·37)	1·22 (1·07–1·38)
Lifestyle			
Alcohol intake frequency			
Never or occasional	1 (ref)	1 (ref)	1 (ref)
1–2 times per week	0·95 (0·87–1·03)	0·95 (0·88–1·03)	0·95 (0·87–1·04)
3–4 times per week	0·89 (0·82–0·97)	0·89 (0·82–0·98)	0·90 (0·82–1·00)
Daily or almost daily	0·98 (0·90–1·07)	0·99 (0·91–1·08)	1·04 (0·95–1·14)
Smoking status			
Never smokers	1 (ref)	1 (ref)	1 (ref)
Current or past occasional smokers	1·23 (1·10–1·36)	1·22 (1·10–1·36)	1·26 (1·13–1·41)
Pack years <10	1·24 (1·09–1·41)	1·22 (1·08–1·39)	1·23 (1·07–1·41)
Pack years 10–19	2·00 (1·80–2·21)	1·94 (1·75–2·16)	1·90 (1·69–2·13)
Pack years ≥20	5·57 (5·18–6·00)	5·51 (5·12–5·94)	5·51 (5·08–5·98)

(Table 3 continues on next page)

	Model 1 (n_1=96 779)	Model 2 (n_2=94 265)	Model 3 (n_3=77 679)
(Continued from previous page)			
Residential tenureship			
Own outright	1 (ref)	1 (ref)	1 (ref)
Mortgage	0·99 (0·92–1·07)	0·99 (0·91–1·07)	0·96 (0·88–1·05)
Rent	1·49 (1·33–1·65)	1·45 (1·30–1·62)	1·51 (1·34–1·71)
Townsend deprivation			
Quintile 1, low	1 (ref)	1 (ref)	1 (ref)
Quintile 2	1·13 (1·03–1·23)	1·13 (1·03–1·24)	1·11 (1·01–1·23)
Quintile 3	1·14 (1·04–1·25)	1·15 (1·05–1·26)	1·12 (1·01–1·23)
Quintile 4	1·11 (1·02–1·22)	1·12 (1·01–1·22)	1·12 (1·01–1·23)
Quintile 5, high	1·22 (1·10–1·35)	1·22 (1·10–1·36)	1·22 (1·10–1·36)
Anthropometrics and comorbidities			
Standing height (cm)	1·02 (1·02–1·03)	1·02 (1·02–1·03)	1·02 (1·02–1·03)
Body-mass index status, kg/m²			
<25	1 (ref)	1 (ref)	1 (ref)
≥25 to <30	0·71 (0·66–0·76)	0·71 (0·66–0·76)	0·72 (0·67–0·78)
≥30	0·64 (0·59–0·70)	0·65 (0·60–0·71)	0·65 (0·60–0·72)
Cardiovascular problems			
None	1 (ref)	1 (ref)	1 (ref)
High blood pressure	1·14 (1·06–1·22)	1·13 (1·06–1·22)	1·12 (1·03–1·21)
Heart attack, angina, or stroke	1·25 (1·07–1·45)	1·24 (1·07–1·45)	1·27 (1·08–1·50)
High blood pressure and heart attack, angina, or stroke	1·39 (1·21–1·60)	1·40 (1·21–1·62)	1·35 (1·15–1·58)
Diabetes status			
No diabetes	1 (ref)	1 (ref)	1 (ref)
Diabetes	1·01 (0·89–1·14)	1·02 (0·89–1·15)	1·02 (0·89–1·17)
Parent's COPD status			
No COPD	1 (ref)	1 (ref)	1 (ref)
One parent with COPD	1·50 (1·40–1·61)	1·50 (1·39–1·61)	1·52 (1·40–1·64)
Both parents with COPD	2·29 (1·83–2·87)	2·31 (1·84–2·90)	2·38 (1·86–3·06)
Haematological biomarkers			
White blood cell counts, 10^9 cells per L*			
Quartile 1 (<5·59)	1 (ref)	1 (ref)	1 (ref)
Quartile 2 (5·59–6·54)	1·14 (1·03–1·26)	1·14 (1·03–1·27)	1·14 (1·02–1·27)
Quartile 3 (6·55–7·70)	1·35 (1·23–1·49)	1·35 (1·22–1·49)	1·33 (1·20–1·48)
Quartile 4 (>7·70)	1·93 (1·75–2·11)	1·91 (1·74–2·10)	1·92 (1·73–2·13)
Neutrophil-to-lymphocyte ratio			
Quartile 1 (<1·66)	1 (ref)	1 (ref)	1 (ref)
Quartile 2 (1·66–2·11*)	0·98 (0·90–1·07)	0·97 (0·89–1·06)	0·95 (0·86–1·05)
Quartile 3 (2·11†–2·72)	1·05 (0·96–1·14)	1·05 (0·96–1·15)	1·05 (0·95–1·15)
Quartile 4 (>2·72)	1·12 (1·03–1·22)	1·12 (1·03–1·22)	1·09 (0·99–1·20)
Eosinophil-to-basophil ratio			
Quartile 1 (<2·17)	1 (ref)	1 (ref)	1 (ref)
Quartile 2 (2·17–4·00)	1·09 (1·00–1·18)	1·09 (1·00–1·18)	1·10 (1·00–1·20)
Quartile 3 (4·01–7·50)	1·02 (0·94–1·11)	1·03 (0·95–1·12)	1·03 (0·94–1·13)
Quartile 4 (>7·50)	1·13 (1·04–1·22)	1·12 (1·03–1·22)	1·14 (1·04–1·26)

Data are ORs (95% CI). Model 1, 2, and 3 represent multiple adjusted models developed for $PM_{2.5}$, urbanicity, and residential greenness, respectively. ORs=odds ratios. COPD=chronic obstructive pulmonary disorder defined as per Global Initiative for Chronic Obstructive Lung Disease stage II plus classification. NDVI=normalised difference vegetation index. *2·112299. †2·11236.

Table 3: **Multiple adjusted ORs for COPD associated with environmental risk factors among participants of the UK Biobank cohort**

participants and environmental exposure variables in the analytic sample and bivariate models of COPD see table 1, 2.

Following the bivariate analyses, the final models were adjusted for demographic characteristics (age, sex, household income, highest qualification, employment

status), lifestyle factors (alcohol intake, smoking, residential tenureship), neighbourhood socioeconomic status (Townsend index), anthropometrics (standing height, BMI status), and comorbidities (cardiovascular problems, diabetes, parental COPD), and haematological biomarkers (white blood cell counts, neutrophil-to-lymphocyte ratio, and eosinophil-to-basophil ratio). Exposure to $PM_{2·5}$ was associated with a higher risk of COPD (OR 1·55, 95% CI 1·14–2·10 per 10 μg/m³ increment; p=0·0047; table 3). Each interquartile increment in urbanicity was associated with higher odds of COPD (1·05, 1·01–1·08 per IQR increment; p=0·011). Exposure to residential greenness had a protective effect, being associated with a lower risk of COPD (0·89, 0·84–0·93 per IQR increment; p<0·0001). Sensitivity analyses using an alternate definition of COPD based on the lower limit of normal criteria produced consistent results, reporting a higher risk of COPD per 10 μg/m³ increments in $PM_{2·5}$ (OR 1·65, 95% CI 1·18–2·30; p=0·0036) as well as per interquartile increment in urbanicity (1·04, 1·00–1·08; p=0·042) and lower odds of COPD per interquartile increase in residential greenness (0·95, 0·90–1·00; p=0·043). In sensitivity analyses for testing the confounding effect of residential greenness on the association between COPD and $PM_{2·5}$ exposure (appendix pp 6–7), the introduction of residential greenness resulted in attenuation of the effect estimates and loss of significance of $PM_{2·5}$ (OR 1·17, 95% CI 0·83–1·66 per 10 μg/m³ increment; p=0·37). The result was similar in the case of urbanicity subsequent to adjustment for residential greenness (1·02, 0·98–1·06 per interquartile increment in urbanicity; p=0·29).

The fitted restricted cubic spline models of the dose–response relationship between COPD prevalence and lung function and exposures are shown in figure 1. A consistent positive association between $PM_{2·5}$ exposure and both COPD prevalence and predicted lung function was observed (figures 1A, B). In the models for residential greenness (figure 1C), the beneficial effects of greenness on COPD prevalence were attenuated beyond the threshold of NDVI 0·21. Piece-wise regression identified significant beneficial associations below the identified point of inflection (OR 0·74, 95% CI 0·67–0·81; p<0·0001), while, above it, the results remained non-significant (0·89, 0·77–1·02; p=0·092). The trend remained consistent in our models of lung function with a non-beneficial effect in the high green areas (figure 1D). Tests for non-linearity of the effects of urbanicity on COPD (p=0·73) and lung function (p=0·15) were however not significant with Harrel's knots. The results of our interaction effects model across quartiles of white blood cell counts indicate that the detrimental effects of $PM_{2·5}$ exposure on lung function were pronounced among participants in the highest white blood cell count quartile in reference to the lowest ($p_{interaction}$=0·0003).

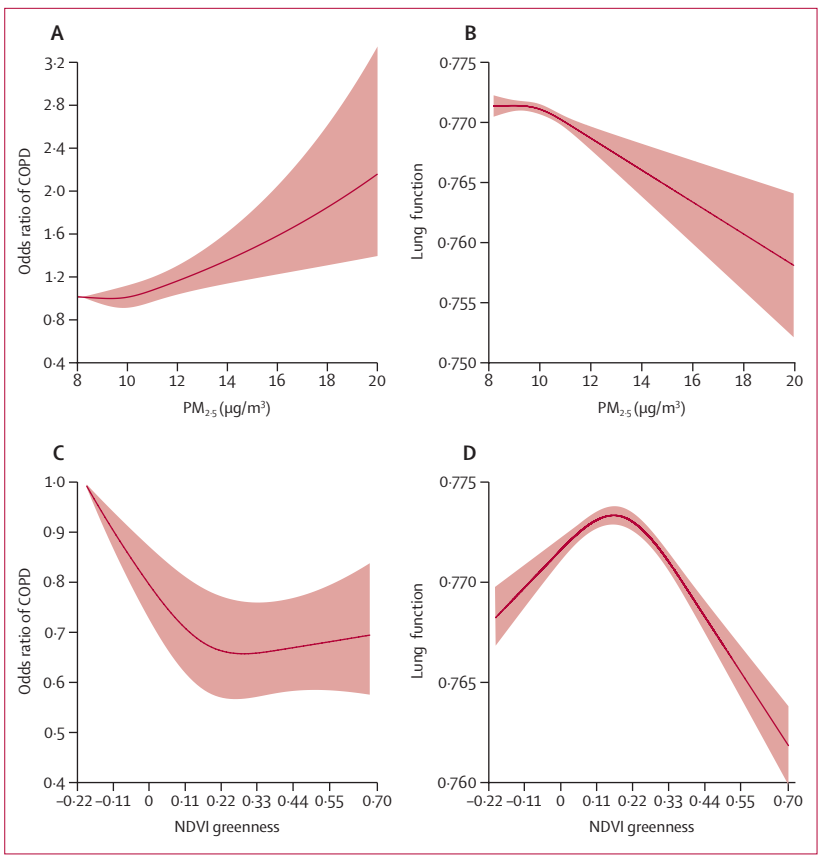

Figure 1: **Association of COPD prevalence and lung function with $PM_{2·5}$ and NDVI greenness, allowing for non-linear effects**

The continuous line represents the estimated odds of COPD and mean lung function and shaded areas represent 95% CIs. Separate models were fitted for odds of COPD and lung function, with restricted cubic splines with Harrell's knots, adjusting for demographic characteristics (age, sex, household income, highest qualification, employment status), lifestyle factors (alcohol intake frequency, smoking, residential tenureship), neighbourhood socioeconomic status (Townsend index), anthropometrics (standing height, body-mass index status), comorbidities (cardiovascular problems, diabetes, parental COPD), and haematological biomarkers (white blood cell counts, neutrophil-to-lymphocyte ratio, and eosinophil-to-basophil ratio). Forced expiratory volume in 1 s to forced vital capacity ratio was used to indicate lung function. COPD=chronic obstructive pulmonary disorder, defined as per Global Initiative for Chronic Obstructive Lung Disease stage II plus classification. NDVI=normalised difference vegetation index.

The odds of COPD were also higher in white blood cell quartile 4 (>7·70); however, the interaction was not significant ($p_{interaction}$=0·072; figure 2A).

On rerunning the models by restricting analyses to non-smokers, the detrimental effects of $PM_{2·5}$ became slightly enhanced (OR 1·80, 95% CI 1·04–3·13 per 10 μg/m³ increment; p=0·037) and so was the protective effect of residential greenness (0·82, 0·75–0·89 per IQR increment; p<0·0001; appendix pp 8–10). Stratifying analyses by urbanicity quartiles (table 4) indicated a more pronounced detrimental effect of $PM_{2·5}$ in the high urbanicity areas, being significant only in the highest quartile (OR 1·85, 95% CI 1·01–3·37 per 10 μg/m³ increment; p=0·046). For residential greenness, the protective effects of greenness slightly attenuated towards the high urbanicity areas (OR 0·85, 95% CI 0·77–0·93; p=0·0007 in the second quartile; 0·87,

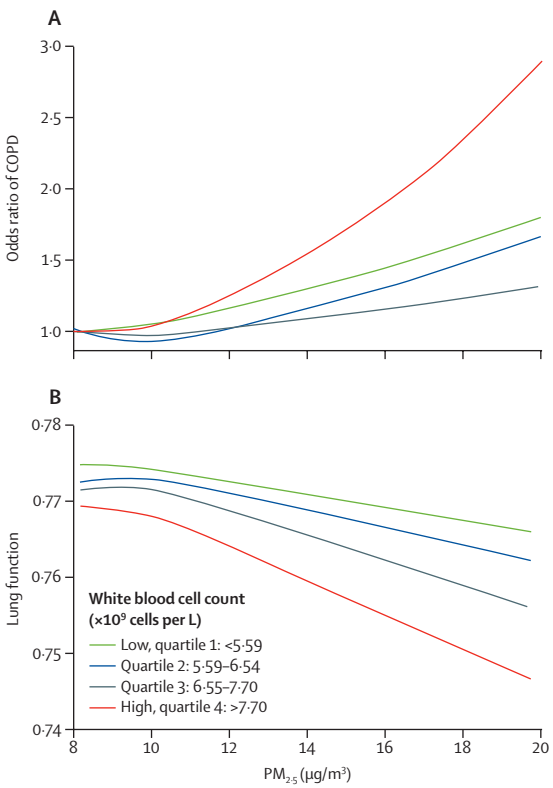

Figure 2: **Associations of COPD and lung function with $PM_{2·5}$ exposure, allowing for effect modification by white blood cell counts**

Separate models were fitted for odds of COPD and lung function, with restricted cubic splines with Harrell's knots, adjusting for demographic characteristics (age, sex, household income, highest qualification, employment status), lifestyle factors (alcohol intake frequency, smoking, residential tenureship), neighbourhood socioeconomic status (Townsend index), anthropometrics (standing height, body-mass index status), comorbidities (cardiovascular problems, diabetes, parental COPD), and haematological biomarkers (neutrophil-to-lymphocyte ratio and eosinophil-to-basophil ratio). Forced expiratory volume in 1 s to forced vital capacity ratio was used to indicate lung function. $p_{interaction}$ in the COPD model=0·0718 (A). $p_{interaction}$ in the lung function model=0·0003 (B). COPD=chronic obstructive pulmonary disorder, defined as per Global Initiative for Chronic Obstructive Lung Disease stage II plus classification.

0·79–0·97; p=0·008 in the third quartile; and 0·89, 0·81–0·98; p=0·019 in the fourth quartile of urbanicity).

Discussion

In this large UK-wide population cohort, we have shown that exposure to $PM_{2·5}$ and urbanicity are associated with higher odds of COPD, whereas residential greenness is associated with reduced odds of COPD, and the results remained consistent after adjustments. This is, to our knowledge, the first large study to examine the effects of $PM_{2·5}$ on COPD and objectively measure attributes of the built environment (high-resolution residential greenness and urbanicity), as well as adjusting for a series of confounders including haematological inflammatory biomarkers. The results remained consistent for an alternative definition of COPD based on the lower limit of normal criteria.

	$PM_{2·5}$ (per 10 µg/m³)	Mean residential greenness (per IQR of NDVI)
Urbanicity quartiles		
Quartile 1, low	1·52 (0·65–3·57)	0·91 (0·83–1·01)
Quartile 2	1·39 (0·60–3·23)	0·85 (0·77–0·93)
Quartile 3	1·87 (0·83–4·21)	0·87 (0·79–0·97)
Quartile 4, high	1·85 (1·01–3·37)	0·89 (0·81–0·98)

Data are odds ratio (95% CI). Models adjusted for demographic characteristics (age, sex, household income, highest qualification, employment status), lifestyle factors (alcohol intake frequency, smoking, residential tenureship), neighbourhood socioeconomic status (Townsend index), anthropometrics (standing height, body-mass index status), comorbidities (cardiovascular problems, diabetes, parental COPD), and haematological biomarkers (white blood cell counts, neutrophil-to-lymphocyte ratio, and eosinophil-to-basophil ratio). COPD=chronic obstructive pulmonary disorder defined as per Global Initiative for Chronic Obstructive Lung Disease stage II plus classification. NDVI=normalised difference vegetation index.

Table 4: **Logistic regression models of association between COPD and environmental factors of $PM_{2·5}$ and residential greenness by urbanicity-quartile subgroups**

Our study reports 54·8% higher odds of COPD per 10 µg/m³ increment in $PM_{2·5}$ exposure. The results are consistent with other previous positive associations between COPD and $PM_{2·5}$ exposure. A 2018 nationwide study of 50 991 participants in China reported twice the odds of COPD (OR 2·00, 95% CI 1·36–2·92) among participants exposed to 75 µg/m³ or more as compared with those in the category of less than 50 µg/m³.[14] A large-scale study reported 39% higher odds of COPD among participants in the highest $PM_{2·5}$ exposure quartile in reference to those in the lowest and associated reductions in lung function.[20] A smaller-scale study of 1872 older adults reported higher odds of COPD (OR 1·21, 95% CI 1·13–1·30) corresponding to 10 µg/m³ increments in $PM_{2·5}$ exposure.[13] Our results remained consistent in the non-linear restricted cubic spline models for COPD prevalence and predicted lung function. Adjusting for the confounding by greenness produced null effects. This finding requires further investigation. As reported previously, it is possible that green spaces can reduce $PM_{2·5}$ loads via absorption and deposition, or dispersion in urban street canyons, as well as provide ventilation corridors breaking air pollution flows.[51,52]

Our study used objective and detailed building footprint-level spatial data to develop an index of urbanicity reporting 4·6% higher odds of COPD per interquartile increment in urbanicity. The detrimental effects of urban areas have been established in terms of rural–urban differences in self-reported COPD diagnosis[14] and the effects of aggregated population density on COPD mortality.[53]

Our study is the first, to our knowledge, to use a very high resolution (0·50 m) measure of residential green exposure and report an overall 11·4% lower odds of COPD per IQR increment in NDVI greenness. A Dutch study of 345 143 participants had previously reported

beneficial effects of 10% increments in green cover within a 1 km residential buffer on COPD and asthma prevalence (OR 0·97, 95% CI 0·96–0·98); however, unlike the present study, prevalence data were derived from routine primary care electronic records and participants' residences were geocoded at the postcode level with green cover expressed as the percentage within 1 km of the postcode centroid in which a participant resided.[29] A UK study reported that green space and gardens were associated with reduced rates of asthma-related hospitalisations,[31] while a Spanish study of children reported beneficial effects on wheezing and bronchitis.[28] Our non-linear restricted cubic spline models, however, indicated that the beneficial effects of greenness on COPD levelled off after a threshold NDVI of 0·21 (figure 2A), while a slight negative effect on lung function was observed beyond this threshold, with a 1·15% net reduction in lung function (figure 2B). We were able to exclude large water bodies in our NDVI calculations and the remaining negative NDVI values correspond to characteristic urban features such as building rooftops and impervious spaces such as roads and parking spaces. High NDVI values are a proxy of low-density semi-urban or rural neighbourhoods with dense vegetation and agricultural land use (appendix pp 3–5). In other words, below the threshold, the observed beneficial effects accrue on account of increasing proportion of green in highly urban areas. Beyond the threshold, highly green areas might potentially increase susceptibility to allergic reactions to pollens[34,54] and affect COPD. Further studies are needed to verify this theory and the potential confounding effects of small water bodies.

That our study detected a slightly pronounced effect of the $PM_{2·5}$ and residential greenness among the non-smoking subgroup is noteworthy from a public health perspective. Approximately 20% of smokers develop COPD; as such, the roles of other factors (ie, environmental, occupational, socioeconomic, nutritional, lung growth, and genetics) are of particular importance. Specific strategies to shield this susceptible subgroup from pollution levels to protect non-smokers from high pollution concentrations warrant further investigation.

With respect to pathophysiological mechanisms, several previous studies have hypothesised pathways via pulmonary or systemic inflammation, establishing direct links between COPD development, reduced lung function endpoints, and the presence of higher concentrations of systemic inflammatory markers.[55] Direct evidence now suggests that exposure to air particulate matter enhances systemic inflammation, thereby causing airflow obstruction, and reducing lung function.[56,57] Our analyses were able to adjust for haematological inflammatory markers, consistently finding higher odds of COPD in the higher quartiles of white blood cell counts, neutrophil-to-lymphocyte ratio, and eosinophil-to-basophil ratio.[46] Consistent with previous studies, interaction effect models in our study indicated that the highest white blood cell quartile constituted a clinically susceptible subgroup, showing reduced lung function and higher odds of COPD with increasing $PM_{2·5}$ exposure concentrations with significant interaction only in the model with lung function. Further longer-term measurements of exposure, lung function, and inflammatory biomarkers are needed to validate this mechanism with confidence.

The observed beneficial effects of residential green exposure on COPD could potentially point to a physical activity-related mechanism, although long-term longitudinal studies are needed to provide support for such hypotheses. We were able to rerun our models accounting for confounding effects of physical activity. Our results remained consistent and physical activity measured as metabolic equivalent of task h per week was beneficially associated with COPD (appendix pp 11–13).

Strengths of the study included the use of high-quality cohort data for the unprecedented size, population, and geographical variability; use of rigorous criteria for assessment of COPD by spirometry; objective and detailed assessment of built environment at a high resolution; and a range of sensitivity tests.

Among limitations, the cross-sectional design constrains causal inference. Our analyses adjusted for several confounders; nonetheless, risk of residual confounding cannot be ruled out as in any observational study. Potential exposure misclassification might have arisen due to the use of ambient $PM_{2·5}$ exposure, which disregards participants' real diurnal activity space-time coordinates and personal exposures profiles could not be disaggregated by indoor or outdoor exposure or time spent commuting. The study did not have $PM_{2·5}$ and built environment data to account for the specific effects of exposures in a participant's work environment or related detrimental occupational exposures. Our assessment of COPD did not involve bronchodilator reversibility testing and hence bronchodilator lung function could not be measured. It has been suggested that reduced lung function on account of airflow obstruction can arise because of COPD or asthma, and so prevalence is likely to be overestimated.[58] To compensate for this overestimation to some extent, we had excluded all COPD cases occurring in conjunction with doctor-diagnosed and self-reported asthma cases (n=709). Additionally, the use of rigorous inclusion criteria (Global Initiative for Chronic Obstructive Lung Disease stage 2+ spirometry) compensated for the absence of post-bronchodilator lung function, ensuring that most COPD cases are likely to have been captured. It has also been suggested that the bronchodilator reversibility test is often unreliable as the test performance can be affected by factors such as the day of testing, the severity of baseline lung-function impairment before testing, and the number of drugs given to test.[59] The UK Biobank cohort had a low response rate of 5·5% and as such the prevalence of COPD and related comorbidities within the cohort are not representative of UK-wide prevalence and our analytical

sample was not representative of the full cohort. The study was also restricted to individuals of white European ancestry. Yet, given the very large sample size and diverse population-level characteristics and heterogeneous environmental exposures across a wide geographical scale, it is probable that the low response rate would have a minimal effect on the generalisability for the reported associations.[60]

In conclusion, in the largest cross-sectional study of COPD prevalence and the built environment thus far, $PM_{2.5}$ exposure and urbanicity were independently associated with higher odds of COPD, whereas green exposure had a net protective effect. Our results suggest the potential value of urban planning and design interventions such as green infrastructure and urban ventilation corridors in minimising or offsetting environmental risks associated with COPD.

Contributors

CS, BZ, and CW conceived the study. CS, BZ, MN, JG, and CW designed the study. CS, SK, and SB helped with the literature review. CS and SK developed the built environment metrics used in the study. CS and BZ did statistical analyses. CS developed the first draft. CW, JG, SB, MN, and SK commented on the draft and all the authors contributed to redrafting and interpretations. All the authors read and approved the final manuscript.

Declaration of interests

We declare no competing interests.

Acknowledgments

The study was funded by the University of Hong Kong's University Research Committee Research Assistant Professorship Grant; UK Biobank seed grant; and UK Economic & Social Research Council's Transformative Research grant (ES/L003201/1). CS acknowledges the National Academy of Medicine–Hong Kong University fellowship in Global Health Leadership. The research used the UK Biobank resource (approved application number 26492). We thank the Ordnance Survey GB, UK's National Mapping Agency, and Manchester Information and Associated Services (University of Manchester) for providing access to its UK-wide spatial data for use in this study.

References

1 Soriano JB, Abajobir AA, Abate KH, et al. Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Respir Med* 2017; **5:** 691–706.

2 Wang H, Naghavi M, Allen C, et al. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016; **388:** 1459–544.

3 Vos T, Barber RM, Bell B, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2015; **386:** 743–800.

4 McLean S, Hoogendoorn M, Hoogenveen RT, et al. Projecting the COPD population and costs in England and Scotland: 2011 to 2030. *Sci Rep* 2016; **6:** 31893.

5 Liu Y, Lee K, Perez-Padilla R, Hudson NL, Mannino DM. Outdoor and indoor air pollution and COPD-related diseases in high- and low-income countries. *Int J Tuberc Lung Dis* 2008; **12:** 115–27.

6 Eisner MD, Anthonisen N, Coultas D, et al. An official American Thoracic Society public policy statement: novel risk factors and the global burden of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2010; **182:** 693–718.

7 Schikowski T, Mills IC, Anderson HR, et al. Ambient air pollution: a cause of COPD? *Eur Respir J* 2014; **43:** 250–63.

8 Berend N. Contribution of air pollution to COPD and small airway dysfunction. *Respirology* 2016; **21:** 237–44.

9 Li J, Sun S, Tang R, et al. Major air pollutants and risk of COPD exacerbations: a systematic review and meta-analysis. *Int J Chron Obstruct Pulmon Dis* 2016; **11:** 3079–91.

10 Cohen AJ, Brauer M, Burnett R, et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *Lancet* 2017; **389:** 1907–18.

11 Ling SH, van Eeden SF. Particulate matter air pollution exposure: role in the development and exacerbation of chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis* 2009; **4:** 233–43.

12 Feng S, Gao D, Liao F, Zhou F, Wang X. The health effects of ambient $PM_{2.5}$ and potential mechanisms. *Ecotoxicol Environ Saf* 2016; **128:** 67–74.

13 Lin H, Qian ZM, Guo Y, et al. The attributable risk of chronic obstructive pulmonary disease due to ambient fine particulate pollution among older adults. *Environ Int* 2018; **113:** 143–48.

14 Wang C, Xu J, Yang L, et al. Prevalence and risk factors of chronic obstructive pulmonary disease in China (the China Pulmonary Health [CPH] study): a national cross-sectional study. *Lancet* 2018; **391:** 1706–17.

15 Liu S, Zhou Y, Liu S, et al. Association between exposure to ambient particulate matter and chronic obstructive pulmonary disease: results from a cross-sectional study in China. *Thorax* 2017; **72:** 788–95.

16 Schikowski T, Adam M, Marcon A, et al. Association of ambient air pollution with the prevalence and incidence of COPD. *Eur Respir J* 2014; **44:** 614–26.

17 Wang Y, Shi L, Lee M, et al. Long-term exposure to PM2. 5 and mortality among older adults in the southeastern US. *Epidemiology* 2017; **28:** 207–14.

18 Pun VC, Kazemiparkouhi F, Manjourides J, Suh HH. Long-term $PM_{2.5}$ exposure and respiratory, cancer, and cardiovascular mortality in older US adults. *Am J Epidemiol* 2017; **186:** 961–69.

19 Rice MB, Ljungman PL, Wilker EH, et al. Long-term exposure to traffic emissions and fine particulate matter and lung function decline in the Framingham Heart study. *Am J Respir Crit Care Med* 2015; **191:** 656–64.

20 Guo C, Zhang Z, Lau AKH, et al. Effect of long-term exposure to fine particulate matter on lung function decline and risk of chronic obstructive pulmonary disease in Taiwan: a longitudinal, cohort study. *Lancet Planet Health* 2018; **2:** e114–25.

21 Ni L, Chuang C-C, Zuo L. Fine particulate matter in acute exacerbation of COPD. *Front Physiol* 2015; **6:** 294.

22 Hooper LG, Young MT, Keller JP, et al. Ambient air pollution and chronic bronchitis in a cohort of U.S. women. *Environ Health Perspect* 2018; **126:** 027005.

23 Cai Y, Schikowski T, Adam M, et al. Cross-sectional associations between air pollution and chronic bronchitis: an ESCAPE meta-analysis across five cohorts. *Thorax* 2014; **69:** 1005–14.

24 Nachman KE, Parker JD. Exposures to fine particulate air pollution and respiratory outcomes in adults using two national datasets: a cross-sectional study. *Environ Health* 2012; **11:** 25.

25 DeVries R, Kriebel D, Sama S. Outdoor air pollution and COPD-related emergency department visits, hospital admissions, and mortality: a meta-analysis. *COPD* 2017; **14:** 113–21.

26 Li M-H, Fan L-C, Mao B, et al. Short-term exposure to ambient fine particulate matter increases hospitalizations and mortality in COPD: a systematic review and meta-analysis. *Chest* 2016; **149:** 447–58.

27 van Dorn A. Urban planning and respiratory health. *Lancet Respir Med* 2017; **5:** 781–82.

28 Tischer C, Gascon M, Fernández-Somoano A, et al. Urban green and grey space in relation to respiratory health in children. *Eur Respir J* 2017; **49:** 1502112.

29 Maas J, Verheij RA, de Vries S, Spreeuwenberg P, Schellevis FG, Groenewegen PP. Morbidity is related to a green living environment. *J Epidemiol Community Health* 2009; **63:** 967–73.

30 Lambert KA, Bowatte G, Tham R, et al. Residential greenness and allergic respiratory diseases in children and adolescents—a systematic review and meta-analysis. *Environ Res* 2017; **159:** 212–21.

31 Alcock I, White M, Cherrie M, et al. Land cover and air pollution are associated with asthma hospitalisations: a cross-sectional study. *Environ Int* 2017; **109:** 29–41.

32 Dadvand P, Villanueva CM, Font-Ribera L, et al. Risks and benefits of green spaces for children: a cross-sectional study of associations with sedentary behavior, obesity, asthma, and allergy. *Environ Health Perspect* 2014; **122:** 1329–35.

33 Lovasi GS, O'Neil-Dunne JP, Lu JW, et al. Urban tree canopy and asthma, wheeze, rhinitis, and allergic sensitization to tree pollen in a New York City birth cohort. *Environ Health Perspect* 2013; **121:** 494–500.

34 Fuertes E, Markevych I, Bowatte G, et al. Residential greenness is differentially associated with childhood allergic rhinitis and aeroallergen sensitization in seven birth cohorts. *Allergy* 2016; **71:** 1461–71.

35 Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015; **12:** e1001779.

36 UK Biobank. Biobank: protocol for a large-scale prospective epidemiological resource. Protocol no: UKBB-PROT-09-06. 2007. http://www.ukbiobank.ac.uk/wp-content/uploads/2011/11/UK-Biobank-Protocol.pdf (accessed Feb 14, 2019).

37 UK Biobank UK. Breath Spirometry. Version 1. 2011. http://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/Spirometry.pdf (accessed Jan 20, 2019).

38 Vogelmeier CF, Criner GJ, Martinez FJ, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease 2017 report. GOLD executive summary. *Am J Respir Crit Care Med* 2017; **195:** 557–82.

39 Eeftens M, Beelen R, de Hoogh K, et al. Development of land use regression models for PM(2·5), PM(2·5) absorbance, PM(10) and PM(coarse) in 20 European study areas; results of the ESCAPE project. *Environ Sci Technol* 2012; **46:** 11195–205.

40 UK Biobank. UK Biobank–Environmental exposures–metadata. 2014. http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/EnviroExposEst.pdf (accessed Jan 24, 2019).

41 Sarkar C, Webster C, Gallacher J. UK Biobank Urban Morphometric Platform (UKBUMP)–a nationwide resource for evidence-based healthy city planning and public health interventions. *Ann GIS* 2015; **21:** 135–48.

42 Markevych I, Schoierer J, Hartig T, et al. Exploring pathways linking greenspace to health: theoretical and methodological guidance. *Environ Res* 2017; **158:** 301–17.

43 Sarkar C. Residential greenness and adiposity: findings from the UK Biobank. *Environ Int* 2017; **106:** 1–10.

44 Sarkar C, Webster C, Gallacher J. Residential greenness and prevalence of major depressive disorders: a cross-sectional, observational, associational study of 94 879 adult UK Biobank participants. *Lancet Planet Health* 2018; **2:** e162–73.

45 Chiaradia AJ, Cooper C, Webster C. sDNA: a software for spatial design network analysis. Specifications, 2012. http://www.cardiff.ac.uk/sdna/ (accessed Feb 14, 2019).

46 Xiong W, Xu M, Zhao Y, Wu X, Pudasaini B, Liu J-M. Can we predict the prognosis of COPD with a routine blood test? *Int J Chron Obstruct Pulmon Dis* 2017; **12:** 615–25.

47 Paliogiannis P, Fois AG, Sotgia S, et al. Neutrophil to lymphocyte ratio and clinical outcomes in COPD: recent evidence and future perspectives. *Eur Respir Rev* 2018; **27:** 170113.

48 Günay E, Sarınç Ulaşlı S, Akar O, et al. Neutrophil-to-lymphocyte ratio in chronic obstructive pulmonary disease: a retrospective study. *Inflammation* 2014; **37:** 374–80.

49 Postma DS, Bush A, van den Berge M. Risk factors and early origins of chronic obstructive pulmonary disease. *Lancet* 2015; **385:** 899–909.

50 Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med* 1999; **159:** 179–87.

51 Jin S, Guo J, Wheeler S, Kan L, Che S. Evaluation of impacts of trees on $PM_{2.5}$ dispersion in urban streets. *Atmos Environ* 2014; **99:** 277–87.

52 Jeanjean APR, Monks PS, Leigh RJ. Modelling the effectiveness of urban trees and grass on $PM_{2.5}$ reduction via dispersion and deposition at a city scale. *Atmos Environ* 2016; **147:** 1–10.

53 Chaix B, Rosvall M, Lynch J, Merlo J. Disentangling contextual effects on cause-specific mortality in a longitudinal 23-year follow-up study: impact of population density or socioeconomic environment? *Int J Epidemiol* 2006; **35:** 633–43.

54 Cariñanos P, Casares-Porcel M. Urban green zones and related pollen allergy: a review. Some guidelines for designing spaces with low allergy impact. *Landsc Urban Plan* 2011; **101:** 205–14.

55 Gan WQ, Man SF, Senthilselvan A, Sin DD. Association between chronic obstructive pulmonary disease and systemic inflammation: a systematic review and a meta-analysis. *Thorax* 2004; **59:** 574–80.

56 Pope CA 3rd, Hansen ML, Long RW, et al. Ambient particulate air pollution, heart rate variability, and blood markers of inflammation in a panel of elderly subjects. *Environ Health Perspect* 2004; **112:** 339–45.

57 Dadvand P, Nieuwenhuijsen MJ, Agustí À, et al. Air pollution and biomarkers of systemic inflammation and tissue repair in COPD patients. *Eur Respir J* 2014; **44:** 603–13.

58 Fortis S, Eberlein M, Georgopoulos D, Comellas AP. Predictive value of prebronchodilator and postbronchodilator spirometry for COPD features and outcomes. *BMJ Open Respir Res* 2017; **4:** e000213.

59 Calverley PM, Albert P, Walker PP. Bronchodilator reversibility in chronic obstructive pulmonary disease: use and limitations. *Lancet Respir Med* 2013; **1:** 564–73.

60 Collins R. What makes UK Biobank special? *Lancet* 2012; **379:** 1173–74.