OPEN ACCESS

ORIGINAL ARTICLE

Long-term exposure to outdoor air pollution and the incidence of chronic obstructive pulmonary disease in a national English cohort

R W Atkinson,[1] I M Carey,[1] A J Kent,[2] T P van Staa,[3,4] H R Anderson,[1,5] D G Cook[1]

Open Access
Scan to access more
free content

ABSTRACT

Objectives The role of outdoor air pollution in the incidence of chronic obstructive pulmonary disease (COPD) remains unclear. We investigated this question using a large, nationally representative cohort based on primary care records linked to hospital admissions.

Methods A cohort of 812 063 patients aged 40–89 years registered with 205 English general practices in 2002 without a COPD diagnosis was followed from 2003 to 2007. First COPD diagnoses recorded either by a general practitioner (GP) or on admission to hospital were identified. Annual average concentrations in 2002 for particulate matter with an aerodynamic diameter <10 μm (PM_{10}) and <2.5 μm ($PM_{2.5}$), nitrogen dioxide (NO_2), ozone and sulfur dioxide (SO_2) at 1 km^2 resolution were estimated from emission-based dispersion models. Hazard ratios (HRs) per interquartile range change were estimated from Cox models adjusting for age, sex, smoking, body mass index and area-level deprivation.

Results 16 034 participants (1.92%) received a COPD diagnosis from their GP and 2910 participants (0.35%) were admitted to hospital for COPD. After adjustment, HRs for GP recorded COPD and PM_{10}, $PM_{2.5}$ and NO_2 were close to unity, positive for SO_2 (HR=1.07 (95% CI 1.03 to 1.11) per 2.2 μg/m^3) and negative for ozone (HR=0.94 (0.89 to 1.00) per 3 μg/m^3). For admissions HRs for $PM_{2.5}$ and NO_2 remained positive (HRs=1.05 (0.98 to 1.13) and 1.06 (0.98 to 1.15) per 1.9 μg/m^3 and 10.7 μg/m^3, respectively).

Conclusions This large population-based cohort study found limited, inconclusive evidence for associations between air pollution and COPD incidence. Further work, utilising improved estimates of air pollution over time and enhanced socioeconomic indicators, is required to clarify the association between air pollution and COPD incidence.

What this paper adds

► Outdoor air pollution is a potential risk factor for the development of chronic obstructive pulmonary disease (COPD).

► Using a large, nationally representative cohort of patients registered with 205 English general practitioners (GP), we investigated the evidence linking exposure to particulate matter and gaseous pollutants to the first diagnosis of COPD recorded either by a GP or on admission to hospital.

► We found limited evidence for associations, although these were comparable with existing cohort evidence based on smaller and less nationally representative populations elsewhere. However, given the ubiquitous nature of the exposure and the substantial burden of COPD on individuals and healthcare services, further work, utilising improved estimates of air pollution exposure for cohort participants, longer follow-up times and more precise disease definition, is required to clarify the role of air pollution in the initiation of COPD.

INTRODUCTION

Cohort studies have demonstrated associations between long-term exposure to outdoor air pollution and adverse health effects.[1] Such studies, exemplified by the American Cancer Society cohort,[2][3] have tended to focus on death as the health outcome with particular emphasis on fine particles and cardiovascular disease.[4] Cohort studies of respiratory mortality have also tended to report positive associations but individually have lacked statistical power.[5] Few studies have focused on chronic obstructive pulmonary disease (COPD).[6] A recent report on the global impact of $PM_{2.5}$ (mass of particulates with aerodynamic diameter <2.5 μm) on COPD was reliant on only three studies, all from the USA and all using mortality as the outcome.[7]

The possibility that air pollution is a risk factor for the initiation of COPD is important given the ubiquitous nature of the exposure and the substantial burden of the disease on individuals and healthcare services. Recent reviews of the literature have been, at the most, only suggestive of a role of ambient air pollution,[6][8] and a recent multicentre cohort study reported limited evidence of associations with air pollution.[9] Few studies have used objective measures of disease[9] or routine databases such as hospital admissions records to identify cases.[10][11]

We have previously used a nationally representative cohort of adults in England to demonstrate positive associations between annual concentrations of air pollution and respiratory mortality including COPD.[5] In this paper, we use a subset of the same cohort to look at the incidence of COPD over time identified from general practitioner (GP) records and from hospital admission records linked with the GP record.

METHODS

The Clinical Practice Research Datalink (CPRD) is a large, validated primary care database that has been collecting anonymous patient data from participating UK general practices since 1987.[12] It includes a full longitudinal medical record for each registered patient containing coded information on diagnoses, prescriptions and tests carried out within the practice. The Hospital Episode Statistics (HES) database routinely records clinical, patient, administrative and geographical information on all National Health Service (NHS)-funded inpatient episodes in the UK. Subject to the practice's approval, the CPRD patient data are routinely linked to HES by a 'trusted third party' via their NHS number, sex, date of birth and postcode and to the Index of Multiple Deprivation (IMD), a small area measure of deprivation used in England for allocation of resources.[13]

We identified 205 English practices, recording high-quality data according to CPRD internal standards (eg, non-contiguous follow-up, incomplete data recording),[14] which had been linked to HES data.[15] From these, we identified 836 557 subjects aged 40–89 years fully registered for at least 1 year on 1/1/2003. Subjects with COPD Read codes on their GP record prior to 2003 (n=24 494) were excluded leaving 812 063 for analysis. First diagnoses of COPD between 2003 and 2007 were identified from: (1) GP records by Read codes indicating a diagnosis of COPD and (2) hospital admissions records by International Classification of Disease codes (ICD-10 Version:2010) (J41–44) indicating a primary diagnosis of COPD. As a sensitivity analysis, we extended the exclusion criteria to include other markers of chronic respiratory disease recorded by the GP including asthma and regular inhaler prescribing (all Read codes available from authors on request). Figure 1 illustrates the cohort selection and exclusions.

Annual mean concentrations in 2002 of particles with a median aerodynamic diameter of $<10\,\mu m$ (PM_{10}), $<2.5\,\mu m$ ($PM_{2.5}$), sulfur dioxide (SO_2), nitrogen dioxide (NO_2) and ozone (O_3) for 1 km grid squares covering England were linked anonymously by the 'trusted third party' from a grid centroid to the nearest residential postcode centroid for each patient. Details of the dispersion models and the model validation including other years are provided in the online supplementary material. Briefly, the models for PM_{10}, $PM_{2.5}$, NO_2 and SO_2 were constructed by estimating emissions from a number of sources (eg, road transport or power generation) before calculating pollution concentrations using a dispersion model. O_3 maps were constructed by interpolating data from rural monitoring stations and adjusting for effects of altitude and NO_x emissions in urban areas.

Model validation was assessed using data from monitoring sites in the national network and from verification sites operated by NETCEN (part of AEA Technology Environment) in conjunction with local authorities and not part of the national network.[16] Model validation was good for NO_2 ($R^2=0.80$ using the national network and 0.57 using verification sites), good for O_3 ($R^2=0.48$ and 0.71), moderate for PM_{10} ($R^2=0.29$ and 0.46) and $PM_{2.5}$ ($R^2=0.23$ and 0.71), but poor for SO_2 ($R^2=0$ and 0.39).

Potential confounders extracted from the GP records included age, sex, smoking status (characterised as non/ex/current, with further categories of 1–19/20–39/40+ cigarettes per day) and body mass index (BMI) (<20, ≥20 and <25, ≥25 and <30, ≥30). The last recorded status before 1/1/2003 was used to code the variables, except for non-smokers, who were reclassified as ex-smokers if they had older historical codes indicating smoking. The amount of current smoking or ex-smoking was estimated by using the maximum cigarettes per day value and grouped accordingly. A 'missing' category was assigned for participants with no recorded smoking status prior to 2003. Individual measures of socioeconomic status (SES) were not available to us and we therefore assessed SES using 2007 IMD deciles, which are available for a small area comprising approximately 1500 people, modified to remove the air quality subdomain.[15] Incidence rates of COPD, calculated from GP and hospital records separately, were compared across all confounders adjusted for age and sex (Stata V.12; StataCorp LP, College Station, Texas, USA).

We used Cox proportional hazards models (SAS V.9.1.3; SAS Institute, Inc, Cary, North Carolina, USA) to investigate associations between pollution concentrations in 2002 and incidence of COPD during 2003–2007. We adjusted cumulatively for (1) age and sex; (2) smoking and BMI and (3) IMD. In a sensitivity analysis, we assessed the impact on hazard ratios (HRs) after adjustment, in turn, for selected individual components of deprivation (income, employment and education). To account for clustering by practice, the modified sandwich estimate of variance was used to produce robust SEs. We further assessed the effect of clustering by fitting a random effect for practice using a shared frailty model in Stata. Two-pollutant models were considered only when the correlation coefficient between pollutants was below 0.5. We also carried out stratified analyses to assess effect modification by the covariates. To allow comparison across pollutants, HRs were quantified for an interquartile range (IQR) change in each pollutant. We also present the main estimates per $10\,\mu g/m^3$ to enable comparison with results of other studies.

RESULTS

Postcode linkage to all pollutants was successful in 99% of patients (table 1). IQR values were lowest for $PM_{2.5}$ ($1.9\,\mu g/m^3$) and highest for NO_2 ($10.7\,\mu g/m^3$). There was significant variation in modelled pollution concentrations by practice region and IMD ($p<0.001$). The highest concentrations of particles and NO_2 were observed in London, whereas the highest concentrations of SO_2 and O_3 occurred in the north and south, respectively. Areas with a lower SES were associated with higher concentrations for all pollutants except O_3. Annual concentrations of PM_{10} and $PM_{2.5}$ were strongly correlated with NO_2 ($r=0.8$), moderately correlated with SO_2 ($r=0.5$) and negatively correlated with O_3 ($r=-0.4$).
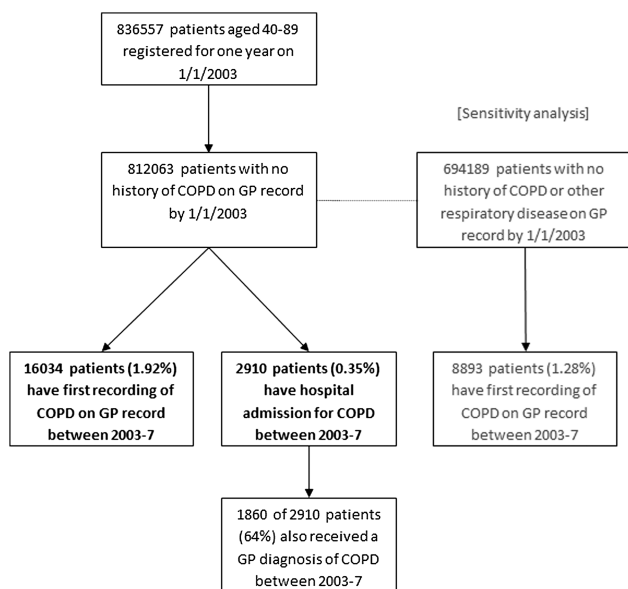
Figure 1 Cohort and case identification.

836557 patients aged 40-89 registered for one year on 1/1/2003

[Sensitivity analysis]

812063 patients with no history of COPD on GP record by 1/1/2003

694189 patients with no history of COPD or other respiratory disease on GP record by 1/1/2003

16034 patients (1.92%) have first recording of COPD on GP record between 2003-7

2910 patients (0.35%) have hospital admission for COPD between 2003-7

8893 patients (1.28%) have first recording of COPD on GP record between 2003-7

1860 of 2910 patients (64%) also received a GP diagnosis of COPD between 2003-7

Table 1 Summary of assigned pollutant concentrations in 2002 for participants with no diagnosis of COPD by end of 2002 (n=812 063)

	Assigned annual average concentration in 2002 (µg/m³)				
	PM_{10}	$PM_{2.5}$	SO_2	NO_2	O_3
No. with pollution linkage (%)	807 401 (99)	807 401 (99)	800 360 (99)	807 000 (99)	801 380 (99)
Mean concentration (SD)	19.7 (2.3)	12.9 (1.4)	3.9 (2.1)	22.5 (7.4)	51.7 (2.4)
IQR	3.0	1.9	2.2	10.7	3.0
Practice region means (SD)					
North [81 practices]	19.8 (2.3)	12.9 (1.5)	4.8 (2.1)	23.4 (6.3)	50.9 (2.4)
South (ex. London) [96 practices]	19.1 (2.0)	12.5 (1.2)	3.2 (1.9)	19.4 (6.1)	52.6 (2.2)
London [28 practices]	22.5 (1.2)	14.6 (0.8)	3.8 (1.2)	33.3 (4.5)	50.2 (0.8)
Test for heterogeneity	p<0.001	p<0.001	p<0.001	p<0.001	p<0.001
Modified IMD quintile means (SD)					
1 (Most deprived)	21.0 (2.2)	13.7 (1.5)	4.7 (2.2)	26.9 (6.3)	50.9 (2.1)
2	21.1 (2.6)	13.8 (1.7)	4.2 (2.3)	26.5 (7.9)	51.3 (2.2)
3	20.4 (2.6)	13.3 (1.7)	4.5 (2.5)	24.2 (8.1)	51.2 (2.2)
4	20.0 (2.3)	13.1 (1.5)	4.4 (2.5)	23.2 (7.5)	51.3 (2.2)
5	19.5 (2.6)	12.7 (1.6)	3.9 (2.4)	21.6 (8.2)	52.1 (2.6)
6	19.3 (2.4)	12.6 (1.5)	3.7 (1.9)	21.1 (7.6)	51.6 (2.6)
7	19.3 (2.3)	12.6 (1.4)	3.9 (2.1)	21.3 (7.5)	51.6 (2.6)
8	19.3 (2.0)	12.6 (1.3)	3.6 (1.8)	20.9 (6.7)	51.8 (2.4)
9	19.4 (1.9)	12.6 (1.2)	3.6 (1.7)	21.8 (6.6)	51.9 (2.3)
10 (Least deprived)	19.6 (1.5)	12.8 (0.9)	3.6 (1.3)	22.1 (5.6)	52.2 (1.9)
Test for trend	p<0.001	p<0.001	p<0.001	p<0.001	p<0.001
Correlation* between pollutants					
$PM_{2.5}$	0.99	–	–	–	–
SO_2	0.45	0.46	–	–	–
NO_2	0.84	0.85	0.38	–	–
O_3	−0.40	−0.39	−0.41	−0.47	–

*Spearman's rank correlation coefficient.
COPD, chronic obstructive pulmonary disease; IMD, Index of Multiple Deprivation; IQR, interquartile range.

Of the 812 063 cohort participants, 16 034 (1.92%) received a COPD diagnosis on their GP record and 2910 (0.35%) were admitted to hospital for COPD during the follow-up period 2003–2007 (figure 1). Of the 2910 patients admitted to hospital with a diagnosis of COPD, 1860 (64%) received a diagnosis of COPD on their GP record by the end of 2007. Extending the exclusion criteria for GP diagnosed COPD to patients with other previously recorded chronic respiratory diseases, such as asthma, reduced the cohort to 694 189 patients, with 8893 (1.28%) patients subsequently receiving an incident diagnosis of COPD.

Table 2 shows the incidence of GP and HES recorded COPD stratified by sex, age, smoking status, BMI, practice region and IMD. Incidence of GP diagnosed COPD was higher in men than in women (2.2% vs 1.8%), but the sex ratio was similar for hospital admissions. Incident cases of COPD increased with age, with only a small reduction in the oldest group (80–89 years) for GP diagnosed COPD. Increasing smoking intensity at baseline, both current and historical, strongly predicted COPD incidence. There was also a strong gradient with IMD, both for GP and hospital recorded COPD, with increasing incidence of both outcomes as deprivation increased. For example, using GP diagnoses, patients in the most deprived decile had an incidence rate almost four times higher than those in the least deprived decile (4.18–1.13%), while for hospital admissions this ratio was almost five (0.88% vs 0.18%). The pattern with IMD was observed across all smoking categories (data not shown). The strong associations with IMD were likely to account for some of the regional variation in COPD incidence, with higher rates in the north compared to the south. However, even among the

most deprived IMD decile, north-south differences were still apparent (4.6% vs 3.4%).

The associations between an IQR change in assigned air pollution concentration in 2002 and a 5-year incidence of COPD recorded on the GP record and for admission to hospital are quantified in a series of HRs in table 3. All pollutants except ozone were positively associated with an increased risk for incident GP recorded COPD when adjusted for age and sex alone. Adjustment for smoking and BMI attenuated all associations, with further adjustment for IMD reducing nearly all associations to the null. The main exception was for SO_2 where a 2.2 µg/m³ increment was associated with an HR=1.07 (95% CI: 1.03 to 1.11). Alternative adjustments for IMD based on individual components (income, education and employment) produced similar reductions in HRs (data not shown). Extending the exclusion criteria to also exclude patients with other previously recorded respiratory disease such as asthma strengthened associations for PM_{10} and $PM_{2.5}$; however, CIs still included unity. In two-pollutant models, the SO_2 association remained robust to adjustment for other pollutants (see online supplementary material table S3). Finally, alternative frailty models using a random effect to account for practice clustering did not materially alter our conclusions (data not shown).

Analyses using incidence based on an admission to hospital for COPD generally produced associations with air pollution greater in magnitude than those seen with GP diagnosis alone (table 3). Again, adjustment for covariates and IMD in particular markedly reduced associations, with all 95% CIs including unity. For example, IQR increments in $PM_{2.5}$ (1.9 µg/m³), SO_2 (2.2 µg/m³) and NO_2 (10.7 µg/m³) were associated with HRs of

Table 2 Age–sex adjusted COPD incidence rates by baseline factors

Baseline variables	Level	All eligible patients N	COPD incidence (GP diagnosed)		COPD incidence (hospital admission)	
			n	Age–sex adjusted (%)	n	Age–sex adjusted (%)
All	–	812 063	16 034	1.92	2910	0.35
Sex	Male	392 574	8382	2.18	1435	0.38
	Female	419 489	7652	1.84	1475	0.36
Age (years)	40–49	241 580	1144	0.47	157	0.07
	50–59	224 817	3472	1.55	400	0.18
	60–69	159 628	4846	3.04	728	0.46
	70–79	119 832	4728	3.95	965	0.81
	80–89	66 206	1844	2.80	660	1.01
Smoking	Non	381 579	2781	0.75	450	0.12
	Ex (unknown)	72 646	1390	1.43	265	0.26
	Ex (1–19 cigarettes/day)	50 670	1906	3.39	334	0.60
	Ex (20–39 cigarettes/day)	23 896	1389	5.63	204	0.90
	Ex (40+ cigarettes/day)	4616	330	6.32	48	0.82
	Current (unknown)	17 048	268	1.28	60	0.26
	Current (1–19 cigarettes/day)	84 386	3449	4.87	685	1.10
	Current (20–39 cigarettes/day)	48 246	2572	7.07	471	1.69
	Current (40+ cigarettes/day)	4246	361	9.20	63	1.73
	Not recorded	124 730	1588	1.27	330	0.29
BMI (kg/m^2)	<20	31 134	935	4.04	232	0.99
	≥20 and <25	262 208	5263	2.16	830	0.35
	≥25 and <30	236 394	4715	1.85	783	0.31
	≥30	105 469	2373	2.23	475	0.46
	Not recorded	176 858	2748	1.72	590	0.36
Practice Region	North	307 826	7270	2.39	1305	0.44
	South (excluding London)	414 539	7259	1.76	1313	0.32
	London	89 698	1505	1.86	292	0.37
IMD Decile	1 (most deprived)	44 037	1785	4.18	380	0.88
	=2	53 439	1640	3.19	323	0.62
	=3	67 133	1906	2.89	371	0.56
	=4	74 147	1726	2.36	354	0.48
	=5	79 579	1586	1.97	287	0.35
	=6	87 057	1648	1.92	293	0.35
	=7	92 102	1517	1.65	289	0.32
	=8	111 891	1737	1.57	251	0.23
	=9	94 877	1361	1.48	195	0.22
	=10 (least deprived)	106 870	1113	1.13	167	0.18

BMI, body mass index; COPD, chronic obstructive pulmonary disease; GP, general practitioner; IMD, Index of Multiple Deprivation.

1.05 (95% CI 0.98 to 1.13), 1.01 (95% CI 0.97 to 1.07) and 1.06 (95% CI 0.98 to 1.15), respectively. The HR for ozone was 0.96 (95% CI 0.90 to 1.02). Restricting the analysis to admissions in participants with COPD confirmed by their GP did not materially alter our findings.

We stratified the adjusted HRs in table 3 by IMD deciles (figure 2). For PM$_{2.5}$ and NO$_2$, there was no consistent evidence of a positive association across different deprivation groups using either GP diagnoses or hospital admission as the outcome. For SO$_2$, the association consistently produced an HR>1 for GP diagnoses, but was less uniform for hospital admissions. The reverse was true for O$_3$ where the HR was generally <1 across the IMD deciles. Stratifying the analyses by other covariates such as age and smoking did not reveal any other patterns (data not shown).

DISCUSSION

In this cohort study of general practice patients, we investigated the relationship between estimates of long-term exposure to outdoor air pollution and the incidence of COPD recorded by GPs and by admission to hospital. While all pollutants except ozone were positively associated with increased risk for incident GP recorded COPD, adjustment for smoking, BMI and IMD in particular reduced associations towards the null except for SO$_2$. Associations with COPD hospital admissions were larger and remained positive for PM$_{2.5}$, PM$_{10}$ and NO$_2$ and negative for O$_3$ after full adjustment, although all CIs included unity. Sensitivity analyses based on excluding asthmatic patients from the analysis, or only counting hospital admissions with a diagnosis confirmed on the GP record, did not alter our findings materially.

Early longitudinal studies of air pollution and the development of COPD used respiratory symptoms to identify participants with COPD[17] [18] while others used lung function measurements, such as forced expiratory volume in the first (FEV$_1$)/forced vital capacity (FVC) ratio[19–21] or the GOLD criterion.[9] A very recent reanalysis and meta-analysis of four European cohorts (in total 6550 participants and including a UK cohort) used FEV$_1$/FVC and the GOLD criterion to define

Table 3 Hazard ratios for the incidence of COPD for an interquartile change in each pollutant

Adjustment factors	PM$_{10}$ (n=807 401)		PM$_{2.5}$ (n=807 401)		SO$_2$ (n=800 360)		NO$_2$ (n=807 000)		O$_3$ (n=801 380)	
	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
Outcome=First diagnosis of COPD on GP record in 2003–2007, Exclusions=Any COPD recorded on GP record prior to 2003										
Age and sex	1.10	1.02 to 1.19	1.12	1.04 to 1.21	1.14	1.09 to 1.19	1.17	1.08 to 1.26	0.88	0.82 to 0.94
+Smoking and BMI	1.06	0.99 to 1.13	1.07	1.00 to 1.15	1.11	1.07 to 1.16	1.11	1.03 to 1.20	0.91	0.85 to 0.97
+IMD	0.99	0.93 to 1.05	1.00	0.94 to 1.06	1.07	1.03 to 1.11	1.03	0.96 to 1.11	0.94	0.89 to 1.00
Full model (10 unit increase)	0.97	0.79 to 1.18	0.97	0.71 to 1.34	1.36	1.14 to 1.62	1.03	0.96 to 1.10	0.82	0.67 to 0.99
Outcome=First diagnosis of COPD on GP record in 2003–2007, Exclusions=Any Markers of Respiratory Disease prior to 2003*										
+IMD	1.01	0.94 to 1.09	1.02	0.94 to 1.09	1.08	1.03 to 1.13	1.04	0.96 to 1.13	0.94	0.87 to 1.00
Outcome=First Hospital Admission for COPD in 2003–2007, Exclusions=Any COPD recorded on GP record prior to 2003										
Age and sex	1.21	1.11 to 1.32	1.23	1.13 to 1.34	1.11	1.05 to 1.17	1.25	1.14 to 1.36	0.87	0.80 to 0.95
+Smoking and BMI	1.15	1.06 to 1.24	1.16	1.08 to 1.26	1.08	1.02 to 1.13	1.17	1.08 to 1.28	0.90	0.84 to 0.97
+IMD	1.05	0.98 to 1.12	1.05	0.98 to 1.13	1.01	0.97 to 1.07	1.06	0.98 to 1.15	0.96	0.90 to 1.02
Full model (10 unit increase)	1.17	0.93 to 1.46	1.31	0.92 to 1.86	1.07	0.86 to 1.33	1.06	0.98 to 1.14	0.87	0.69 to 1.08
Outcome=First Hospital Admission for COPD in 2003–2007 with additional GP diagnosis, Exclusions=Any COPD recorded on GP record prior to 2003										
+IMD	1.05	0.96 to 1.14	1.05	0.97 to 1.14	1.04	0.98 to 1.10	1.04	0.94 to 1.15	0.97	0.90 to 1.05

Hazard ratios are for IQR change in pollutant level unless stated (PM$_{10}$=3.0 µg/m³, PM$_{2.5}$=1.9 µg/m³, SO$_2$=2.2 µg/m³, NO$_2$=10.7 µg/m³, O$_3$=3.0 µg/m³). Practice is accounted for in the above models by the sandwich estimator to produce robust standard errors.
*The numbers of patients in the analyses with this additional exclusion were: PM$_{10}$/PM$_{2.5}$ n=690 102, SO$_2$ n=684 261, NO$_2$ n=689 767, O$_3$ n=684 261.
BMI, body mass index; COPD, chronic obstructive pulmonary disease; GP, general practitioner; HES, hospital episode statistics; IMD, Index of Multiple Deprivation.

COPD and found positive, but not statistically significant, associations with NO$_2$ and PM$_{10}$.[9] Epidemiological studies designed to identify the small associations between air pollution and health outcomes require disease recording in large populations, which precludes the use of individual lung function measurements. Only a small number of studies have used routine administrative databases such as hospital admissions records to identify COPD cases. Andersen et al[10] studied 57 000 participants in a Danish cohort recording their first admission to

hospital for COPD between 1993 and 2006. COPD incidence was associated with 35-year estimated residential mean NO$_2$ concentrations (HR=1.08; 95% CI, 1.02 to 1.14, per IQR of 5.8 µg/m³)—an association larger for the same increment in our study (HR=1.03). Gan et al[11] also studied hospitalisations for COPD in a cohort of almost 500 000 residents in metropolitan Vancouver aged 45–85 years and without a previous physician diagnosis of COPD at baseline. Residential exposures to NO$_2$ estimated from land use regression models were not associated
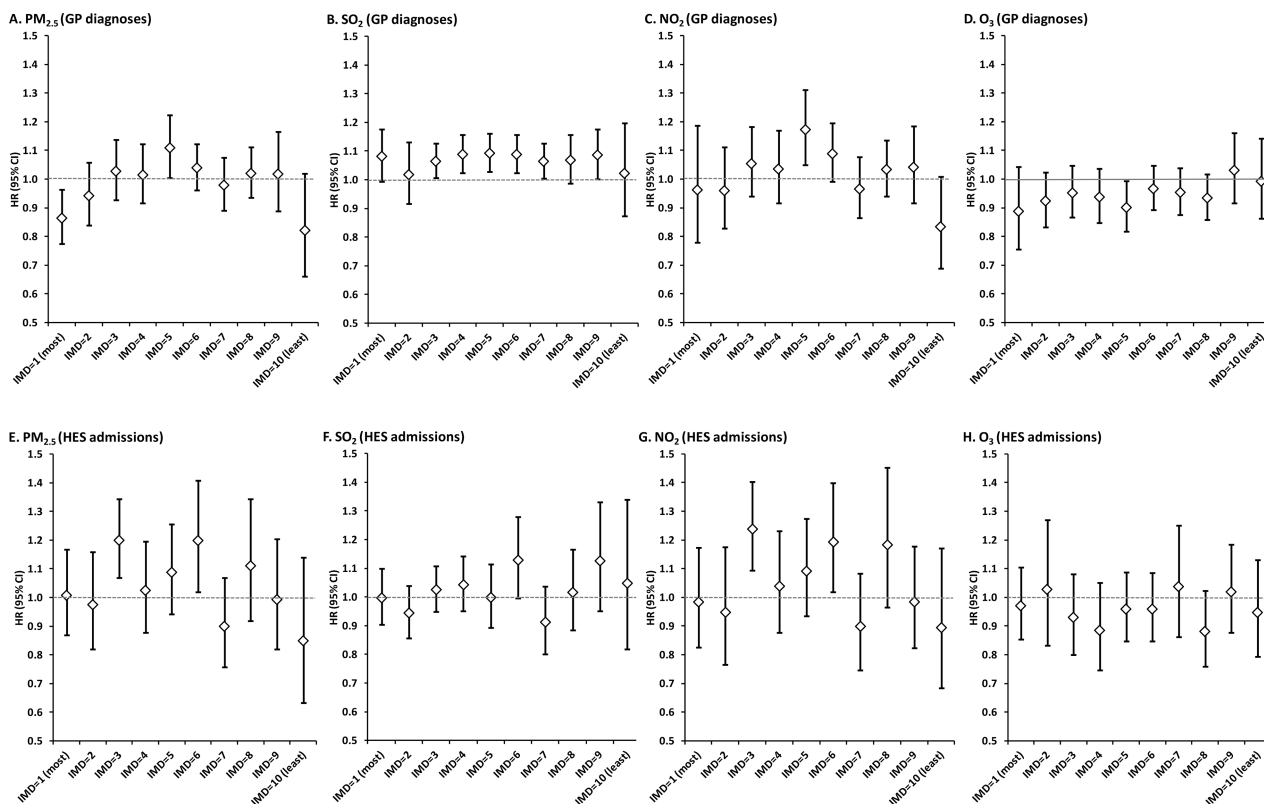

Figure 2 Hazard ratios and 95% confidence intervals for the incidence of Chronic Obstructive Pulmonary Disease for an interquartile change in each pollutant stratified by Index of Multiple Deprivation deciles.

with COPD hospitalisations (HR=1.00). They also found no evidence of an association with $PM_{2.5}$ (HR=1.02 (95% CI 0.98 to 1.06) per 1.58 $\mu g/m^3$ increment, which compares to HR=1.04 for the same increment in our study.

Our finding of an association between long-term exposure to SO_2 and COPD is, to the best of our knowledge, new, although we note that this finding is specific to GP recorded COPD and hospital admission when confirmed by the GP. SO_2 has been associated with respiratory mortality in some, but not all, cohort studies.[5] Short-term exposure time-series studies have indicated associations between daily concentrations of SO_2 and emergency hospital admissions for COPD[22] but not GP consultations for lower respiratory disease excluding asthma.[23] Given the lack of a plausible hypothesis linking current low levels of SO_2 to health,[24] it is possible that our finding reflects historical exposure to higher concentrations of SO_2, or that our SO_2 model, which showed an association with cardiovascular disease in a previous analysis,[15] is providing a geographical proxy of poor overall health or acts as a marker for more toxic substances produced by the combustion of sulfur-containing fossil fuel.

We also found that associations with ozone were consistently negative, which we have seen previously with cardiovascular disease and mortality.[5 15] While this is in part explained by negative correlations between ozone and the other pollutants, the smaller variation of ozone concentration by area probably results in it better representing regional levels. An analysis confined to the months in which ozone levels are highest might have been informative, but we were unable to stratify by season.[2]

In our study, we used two distinct data sources to identify COPD over time—GP diagnoses and hospital admissions. The use of a nationally representative UK primary care database to study COPD incidence has been established,[25] but some important changes that might affect recording have taken place. One of the most notable changes was the introduction of the Quality and Outcomes Framework in 2004 which required GPs to confirm their COPD diagnoses by spirometry from 2006 onwards. This change may have temporarily introduced additional variability in the recording of diagnoses across practices. However, the introduction of spirometry should improve diagnostic certainty and lead to improved and more specific outcome measures for use in epidemiological studies of air pollution where large sample sizes are required to detect relatively small effects.

We also used hospital admissions for COPD as an outcome. Andersen et al[10] noted the appeal of using an objective measure such as first-ever hospital admission assessed objectively from a nationwide register but recognised that such patients represented only those with severe or poorly controlled COPD. Hence, admissions are unlikely to represent the real burden of COPD in the population.[10] In this respect, our study had the benefit of two large, quality assured, independent sources of data to identify incidence: hospital admissions records to identify participants at the severe end of the spectrum of disease and primary care records incorporating participants with COPD who had not yet required admission to hospital. However, the consistency in recording of a COPD diagnosis between the two was not complete—36% of incident hospital admissions for COPD were not confirmed by a corresponding GP diagnosis during the study. A sensitivity analysis restricting hospital admissions to those corroborated by a subsequent GP diagnosis of COPD did not alter substantively the associations observed with hospital admissions. While the use of large-scale databases allows us to detect small effects in large populations, the lack of

consistency between the two outcomes limits the conclusions that can be drawn from each analysis.

We found that adjustment for deprivation had a substantial effect on all HRs, moving all associations towards unity and statistical non-significance. The direction of this adjustment is not surprising as lower SES has been shown to be related to poorer air quality and respiratory function in England previously.[26] Adjustment for deprivation is important as it is associated with air pollution as well as other factors that might be correlated with COPD outcomes.[27] We were limited by relying on census measures of deprivation based on a geographical area of approximately 1500 people;[13] however, neighbourhood SES indicators can be an acceptable proxy for individual measures.[28] It is possible that in urban areas the IMD may represent fine-scale variability in actual pollution levels not represented by our modelled estimates. For example, within the modelled 1×1 km area, patients living closer to busy roads are assigned identical exposure as those further away, but may have different deprivation profiles.[29] While this may raise the possibility that we have overadjusted in our models, when we stratified our analysis by IMD deciles, we found little evidence of consistent effects of air pollution within the separate groups and no evidence of greater effects in areas of low socioeconomic characteristics.[30] It is also possible that some of the IMD adjustment may be acting as a proxy for further smoking history, as our smoking variable was unable to calculate lifetime pack years.

We have discussed previously the performance and validation (see online supplementary material) of our air pollution dispersion models.[5] We note that the validity of the modelled exposure data (as measured by R^2 statistics in relation to monitored data) varied among the pollutants and from year to year. In 2002, our chosen year for exposure assignment, model validation was good for NO_2, moderate for PM_{10} and poor for SO_2, although this varied substantially from year to year (R^2=0.23–0.45 at national network sites and 0–0.6 at the verification sites). Models for ozone demonstrated good to reasonable model performance. The difficulty in modelling PM and SO_2 has been well documented[31 32] and we note also that alternative pollution models have reported similar performance when applied to the UK.[31 32] Other sensitivity analysis using exposure estimates averaged over a number of years, which include better validation statistics, largely confirmed our findings based on the 2002 data only including associations with SO_2 (data not shown). Nonetheless, we believe that the poorer model performance for PM_{10} and SO_2 relative to NO_2 should be taken into consideration in the overall assessment of our results.

In comparison with other cohort studies of COPD, our pollution estimates differed in terms of spatial resolution (1×1 km grids vs residential,[10] postcode[11]) and historical exposure (35[10] and 5 years[11]). The question of historical versus recent exposure has been considered previously[33] in relation to mortality in the American Cancer Society cohort[3] and no strong evidence was found to support earlier over more recent pollution exposure estimates. Annual estimates of pollution concentrations from our model for other years around 2002 suggested strong correlations over time, so a single exposure year provides an acceptable proxy of longer exposure over time. However, we note that our estimates of pollution exposure may not adequately represent cumulative exposure over prior decades, especially where concentrations have generally been falling over time. A further limitation of our study was the relatively short follow-up period (up to 5 years), although this was mitigated to some extent by the size of the study population and number of events (>16 000

and >2900 incident COPD cases recorded by a GP and admission, respectively).

In a previous analysis of the same cohort,[5] we found associations with COPD and respiratory mortality, which were comparable with those for hospitalisations reported in this case; however, owing to the relatively smaller number of mortality outcomes (less than 30%), the precision of our estimates was reduced. Hospital admissions may not represent a true incident point but may instead indicate the role of air pollution in exacerbating COPD. This hypothesis is supported by evidence from short-term exposure time-series studies which have found positive associations between daily concentrations of air pollution to increase in numbers of COPD hospital admissions and mortality in urban populations.[22 34 35]

In conclusion, this large population-based cohort study found only limited, inconclusive evidence for associations between air pollution concentrations and COPD incidence and inconsistency between results for COPD identified from GP records and from hospital admissions. Given the ubiquitous nature of the exposure and the substantial social and economic burden of the disease, further work utilising improved estimates of air pollution exposure for cohort participants, individual markers of SES, longer follow-up times and a more precise disease definition is required to clarify the role of air pollution in the initiation of COPD.

REFERENCES

1 Hoek G, Krishnan RM, Beelen R, et al. Long-term air pollution exposure and cardio-respiratory mortality: a review. *Environ Health* 2013;12:43.
2 Krewski D, Jerrett M, Burnett RT, et al. Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality. *Res Rep Health Eff Inst* 2009;140:5–114.
3 Pope CA III, Burnett RT, Thun MJ, et al. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA* 2002;287:1132–41.
4 Brook RD, Rajagopalan S, Pope CA III, et al. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. *Circulation* 2010;121:2331–78.
5 Carey IM, Atkinson RW, Kent AJ, et al. Mortality associations with long-term exposure to outdoor air pollution in a national english cohort. *Am J Respir Crit Care Med* 2013;187:1226–33.
6 Schikowski T, Mills IC, Anderson HR, et al. Ambient air pollution: a cause of COPD? *Eur Respir J* 2014;43:250–63.
7 Lim SS, Vos T, Flaxman AD, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012;380:2224–60.
8 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.
9 Schikowski T, Adam M, Marcon A, et al. Association of ambient air pollution with the prevalence and incidence of COPD. *Eur Respir J*. 31 Jan 2014. Epub ahead of print.
10 Andersen ZJ, Hvidberg M, Jensen SS, et al. Chronic obstructive pulmonary disease and long-term exposure to traffic-related air pollution: a cohort study. *Am J Respir Crit Care Med* 2011;183:455–61.
11 Gan WQ, FitzGerald JM, Carlsten C, et al. Associations of ambient air pollution with chronic obstructive pulmonary disease hospitalization and mortality. *Am J Respir Crit Care Med* 2013;187:721–7.
12 Medicines and Healthcare Regulatory Agency. The Clinical Practice Research Datalink. 2012. http://www.cprd.com (accessed 1-4-2014)
13 Noble M, McLennan D, Wilkinson K, et al. *The English indices of deprivation 2007*. Department for Communities and Local Government, 2008.
14 Williams T, van Staa TP, Puri S, et al. Recent advances in the utility and use of the General Practice Research Database as an example of a UK Primary Care Data resource. *Ther Adv Drug Saf* 2012;3:89–99.
15 Atkinson RW, Carey IM, Kent AJ, et al. Long-term exposure to outdoor air pollution and incidence of cardiovascular diseases. *Epidemiology* 2013;24:44–53.
16 Stedman JR, Handley C. *A comparison of national maps of NO2 and PM10 concentrations with data from the NETCEN 'Calibration Club.'* Report AEAT/ENV/R/0725. AEA Technology, 2001.
17 Euler GL, Abbey DE, Magie AR, et al. Chronic obstructive pulmonary disease symptom effects of long-term cumulative exposure to ambient levels of total suspended particulates and sulfur dioxide in California Seventh-Day Adventist residents. *Arch Environ Health* 1987;42:213–22.
18 Euler GL, Abbey DE, Hodgkin JE, et al. Chronic obstructive pulmonary disease symptom effects of long-term cumulative exposure to ambient levels of total oxidants and nitrogen dioxide in California Seventh-Day Adventist residents. *Arch Environ Health* 1988;43:279–85.
19 Abbey DE, Burchette RJ, Knutsen SF, et al. Long-term particulate and other air pollutants and lung function in nonsmokers. *Am J Respir Crit Care Med* 1998;158:289–98.
20 Downs SH, Schindler C, Liu LJ, et al. Reduced exposure to PM10 and attenuated age-related decline in lung function. *N Engl J Med* 2007;357:2338–47.
21 Gotschi T, Sunyer J, Chinn S, et al. Air pollution and lung function in the European Community Respiratory Health Survey. *Int J Epidemiol* 2008;37:1349–58.
22 Department of Health. Quantitative systematic review of short term associations between ambient air pollution (particulate matter, ozone, nitrogen dioxide, sulphur dioxide and carbon monoxide), and mortality and morbidity. St George's, University of London, 2007.
23 Hajat S, Haines A, Goubet SA, et al. Association of air pollution with daily GP consultations for asthma and other lower respiratory conditions in London. *Thorax* 1999;54:597–605.
24 Department of Health. Long-term exposure to air pollution: effect on mortality. Committee on the Medical Effects of Air Pollutants, 2009.
25 Simpson CR, Hippisley-Cox J, Sheikh A. Trends in the epidemiology of chronic obstructive pulmonary disease in England: a national study of 51 804 patients. *Br J Gen Pract* 2010;60:277–84.
26 Wheeler BW, Ben-Shlomo Y. Environmental equity, air quality, socioeconomic status, and respiratory health: a linkage analysis of routine data from the Health Survey for England. *J Epidemiol Community Health* 2005;59:948–54.
27 Gershon AS, Dolmage TE, Stephenson A, et al. Chronic obstructive pulmonary disease and socioeconomic status: a systematic review. *COPD* 2012;9:216–26.
28 Adams J, Ryan V, White M. How accurate are Townsend Deprivation Scores as predictors of self-reported health? A comparison with individual level data. *J Public Health (Oxf)* 2005;27:101–6.
29 Cesaroni G, Badaloni C, Romano V, et al. Socioeconomic position and health status of people who live near busy roads: the Rome Longitudinal Study (RoLS). *Environ Health* 2010;9:41.
30 Jerrett M, Burnett RT, Brook J, et al. Do socioeconomic characteristics modify the short term association between air pollution and mortality? Evidence from a zonal time series in Hamilton, Canada. *J Epidemiol Community Health* 2004;58:31–40.
31 Vienneau D, de Hoogh K, Beelen R, et al. Comparison of land-use regression models between Great Britain and the Netherlands. *Atmos Environ* 2010;44:688–96.
32 Gulliver J, Morris C, Lee K, et al. Land use regression modeling to estimate historic (1962–1991) concentrations of black smoke and sulfur dioxide for Great Britain. *Environ Sci Technol* 2011;45:3526–32.
33 Department of Health. The Mortality Effects of Long-Term Exposure to Particulate Air Pollution in the United Kingdom. Committee on the Medical Effects of Air Pollutants, 2010.
34 Sunyer J. Urban air pollution and chronic obstructive pulmonary disease: a review. *Eur Respir J* 2001;17:1024–33.
35 Anderson HR, Spix C, Medina S, et al. Air pollution and daily admissions for chronic obstructive pulmonary disease in 6 European cities: results from the APHEA project. *Eur Respir J* 1997;10:1064–71.