Long-Term Exposure to Outdoor Air Pollution and Incidence of Cardiovascular Diseases

Richard W. Atkinson,† Iain M. Carey,† Andrew J. Kent,‡ Tjeerd P. van Staa,§ H. Ross Anderson,¶ and Derek G. Cook†

Background: Evidence based largely on US cohorts suggests that long-term exposure to fine particulate matter is associated with cardiovascular mortality. There is less evidence for other pollutants and for cardiovascular morbidity. By using a cohort of 836,557 patients age 40 to 89 years registered with 205 English general practices in 2003, we investigated relationships between ambient outdoor air pollution and incident myocardial infarction, stroke, arrhythmia, and heart failure over a 5-year period.

Methods: Events were identified from primary care records, hospital admissions, and death certificates. Annual average concentrations in 2002 for particulate matter with a median aerodynamic diameter <10 (PM_{10}) and <2.5 microns, nitrogen dioxide (NO_{2}), ozone, and sulfur dioxide at a 1 × 1 km resolution were derived from emission-based models and linked to residential postcode. Analyses were performed using Cox proportional hazards models adjusting for relevant confounders, including social and economic deprivation and smoking.

Results: While evidence was weak for relationships with myocardial infarction, stroke, or arrhythmia, we found consistent associations between pollutant concentrations and incident cases of heart failure. An interquartile range change in PM_{10} and in NO_{2} (3.0 and 10.7 μg/m³, respectively) both produced a hazard ratio of 1.06 (95% confidence interval = 1.01–1.11) after adjustment for confounders. There was some evidence that these effects were greater in more affluent areas.

Conclusions: This study of an English national cohort found evidence linking long-term exposure to particulate matter and NO_{2} with the development of heart failure. We did not, however, replicate associations for other cardiovascular outcomes that have been reported elsewhere.

(Epidemiology 2013;24: 44–53)

Exposure to outdoor particulate air pollution is believed to be associated with cardiovascular disease (CVD) mortality and morbidity.1,2 Mechanistic studies have suggested a number of potential pathways from exposure to disease, including release of proinflammatory mediators, changes to the systemic autonomic nervous system balance, and translocation of particles or particle constituents into the circulation.2 Epidemiologic evidence for adverse health effects comes from studies of exposure over short (hours to days) and long (eg, a few years) periods. Evidence for the latter is provided by cohort studies, most of which have focused on all-cause and cardiopulmonary mortality. Exposure estimates for cohort members have been derived from average concentrations of fine particles derived from fixed-site community monitors3–10 or have used spatial modeling techniques to estimate exposure on a finer spatial scale.11–17 Studies of CVD mortality report positive associations with fine particles.10,18–21 Only a few studies have included nonfatal outcomes (such as hospitalization) in their identification of incident events.13–16,20 Furthermore, only a small number of studies have considered other regulated pollutants such as nitrogen dioxide (NO_{2}), ozone, and sulfur dioxide (SO_{2}) in relation to CVD with less consistent results.22,23

We investigated the associations between long-term exposure to a number of ambient outdoor air pollutants and the incidence of a range of CVD events using a national cohort of English adults registered with their general practitioner. Linkage of primary care records to national registers of hospital admissions and mortality records, together with small-area measures of social and economic deprivation, maximized case identification and control for important confounders. Annual concentrations of a range of pollutants at a 1 × 1 km grid spatial resolution were derived from national models that incorporate distant, point, and local sources.24

Submitted 25 October 2011; accepted 27 July 2012.

From the †Division of Population Health Sciences and Education and MRC-HPA Centre for Environment and Health, St George’s, University of London, London, United Kingdom; ‡AEA Technology plc, Harwell IBC, Didcot, Oxfordshire, United Kingdom; §Clinical Practice Research Data-link, Medicines and Healthcare Products Regulatory Agency, London, United Kingdom; and ¶MRC-HPA Centre for Environment and Health, King’s College London, London, United Kingdom. Supported by funding from Policy Research Programme in the Department of Health.

The authors report no conflict of interests.

Copyright © 2012 by Lippincott Williams & Wilkins
ISSN: 1044-3983/13/2401-0044
DOI: 10.1097/EDE.0b013e318276ccb8
METHODS

Data Sources

Clinical Practice Research Datalink

The Clinical Practice Research Datalink (formerly the General Practice Research Database) is a large validated primary care database that has been collecting anonymized patient data from participating UK general (family) practices since 1987. As of 2010, it represents approximately 8% of the UK population of which 98% is believed to be registered with a general practitioner. It includes a full longitudinal medical record for each patient, with information on diagnoses, prescriptions, and tests performed within the practice and information from outside sources such as outpatient consultations and admissions to hospital.

Hospital Episode Statistics and Death Registration Data

Details on every hospital admission in the United Kingdom were provided by the Hospital Episode Statistics database. This system records clinical, patient, administrative, and geographical information on all National Health Service (NHS) funded inpatient episodes. Death registration data are processed by the Office for National Statistics in the United Kingdom. Details recorded include age, sex, address, and both the immediate and underlying causes of death coded according to the International Classification of Diseases (ICD). Subject to the practice’s approval, the Clinical Practice Research Datalink patients are linked to both Hospital Episode Statistics and Office for National Statistics data by a “trusted third party” using their National Health Service (NHS) number, sex, date of birth, and postcode. As of 2010, this linkage has been performed for 224 Clinical Practice Research Datalink practices (approximately 40% of all practices).

Study Group

Eligibility

We selected 205 English practices that, by 1 January 2003, (1) were recording data that were deemed to be “Up-to-Standard” (an internal Clinical Practice Research Datalink quality control metric indicating the date from which a practice was continually recording high-quality data), and (2) had linked hospital admission and mortality data available from this date. From these practices, we identified a cohort of patients who were between the age of 40 and 89 years in 2003, were fully registered on 1 January 2003, and had been continually so for 12 months. This resulted in 836,557 patients eligible for our study.

Outcomes

An incident event was defined as the first event occurring between 2003 and 2007 recorded on any of our three data sources: Clinical Practice Research Datalink (Read codes), Hospital Episode Statistics (ICD codes), and Office for National Statistics mortality (ICD codes). The main conditions and their ICD-10 codes were coronary heart disease (CHD), I20-25; myocardial infarction (MI), I21-23; cerebrovascular disease, I60-69; stroke, I61, I63-64; arrhythmias, including cardiac arrest, I46-49, R00.1; and heart failure, I50. These were manually mapped to corresponding Read codes (list available from authors). Patients with disease recorded before 1 January 2003 (or with no date attached to the code) were excluded from the analysis for that outcome. Thus, a patient with angina first recorded in 2000 would still be eligible for analyses of MI and other outcomes.

Air Pollution

Annual mean concentrations of particles with a median aerodynamic diameter of <10 µm (particulate matter, PM$_{10}$) and <2.5 µm (PM$_{2.5}$), SO$_2$, NO$_2$, and ozone (O$_3$) for 1 × 1 km grids covering England were estimated using air dispersion models. Details of the modeling methodology are given in the eAppendix (http://links.lww.com/EDE/A628). In brief, we constructed the models for PM$_{10}$, PM$_{2.5}$, NO$_2$, and SO$_2$ by identifying all known emission sources by emission sector (eg, power generation, domestic combustion, road traffic, industry, waste) and estimating quantities of emissions. Pollution concentrations were calculated by summing the estimated concentrations for pollutant-specific components from distant sources (characterized by rural background concentrations, interpolated from rural measurements), point sources (calculated using an air dispersion model), and local area sources (calculated using a kernel-based air dispersion model) including the effect of weather conditions. 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 by using the concentrations calculated with the air dispersion models.

Model validation was assessed using data from the national air quality monitoring networks (also used in the global calibration of the model) and from a separate network of monitors also yielding high-quality data (verification sites). Further details are given in the eAppendix (http://links.lww.com/EDE/A628). In brief, model validation was good for NO$_2$ (eg, in 2002, the $R^2$ statistics were 0.80 for the national network sites and 0.57 for the verification sites). For PM$_{10}$, there was moderate agreement ($R^2 = 0.29$ for national network sites and $R^2 = 0.46$ for verification sites in 2002). Model performance statistics for PM$_{2.5}$ were not available for 2002 (because of the paucity of monitoring data available at that time), but in later years, the performance was good (eg, in 2008, $R^2$ statistic of 0.5 at six verification sites). Concentrations of PM$_{10}$ and PM$_{2.5}$ were highly correlated ($r > 0.9$). Therefore, to simplify the presentation of our findings, we report results for PM$_{10}$ in the article and present results for PM$_{2.5}$ in the eTables (http://links.lww.com/EDE/A628). For SO$_2$, however, the validation was moderate to poor and varied substantially from year to year (from 2002 to 2006, the $R^2$ statistics varied from 0.23 to 0.45 at national network sites and 0 to 0.6 at the verification...
Data Linkage

Annual pollution concentrations for each 1 × 1 km grid were linked to the patients’ residential postcodes in the Clinical Practice Research Datalink by a “trusted third party” to ensure anonymity. In England, the 1.4 million active postcodes vary geographically in size but identify on average about 15 residential addresses. The centroid of each postcode was mapped to the nearest centroid of each 1 × 1 km grid. Across the 205 practices, the number of 1 × 1 km grids with a registered patient ranged from 27 to 252.

Covariates

We extracted the following patient information from the Clinical Practice Research Datalink to be used as covariates in our models: age, sex, smoking, body mass index (BMI), diabetes, and hypertension. All covariates had to be recorded by the start of follow-up (1 January 2003). Practice location was designated as being in the North or South or in Greater London.

Socioeconomic status was assessed by using the Index of Multiple Deprivation 2007, a composite small-area (approximately 1500 people) measure of deprivation used in England for allocation of resources. The published index was constructed by a weighted score of seven domains: income, employment, health deprivation and disability, education skills and training, barriers to housing and services, crime, and living environment. Because the living environment domain contains a subdomain relating to air quality, we recalculated the overall index excluding this domain. We then reranked areas nationally and summarized them as deciles.

Statistical Analyses

We used Cox proportional hazards models (PROC PHREG in SAS version 9.1.3; SAS Institute, Cary, NC) to investigate the association between outcome and air pollution. Time to event was modeled in months, and patients were censored if they deregistered from the practice (or died) without experiencing an event before 31 December 2007. We considered two approaches to modeling pollution: a fixed level (2002) and a varying time-dependent covariate (based on previous years’ exposure). In the models, we adjusted cumulatively for the following: (1) age and sex; (2) smoking, BMI, diabetes, and hypertension; and (3) Index of Multiple Deprivation. To allow comparison across pollutants, hazard ratios (HRs) were calculated for an interquartile range (IQR) change in each pollutant (given in Table 1). Air pollutants were modeled individually and in pairs in two-pollutant models.

To investigate the clustering impact of air pollution levels within practice, we calculated intraclass correlation coefficients. In the Cox models, we accounted for practice in two ways. First, we used a robust sandwich estimate for the covariance matrix, which results in robust standard errors for the effect estimates. Second, we added to the full model a term for “practice mean exposure,” which allows us to derive the contribution of between- and within-exposure to the overall effect.

RESULTS

Air Pollution Linkage

Successful postcode linkage was made for approximately 99% of patients in our cohort for all pollutants of interest (Table 1). Annual mean concentrations of PM\textsubscript{10} and NO\textsubscript{2} were higher in London than the rest of the country, whereas SO\textsubscript{2} was highest in the North. PM\textsubscript{10}, SO\textsubscript{2}, and NO\textsubscript{2} were all associated with greater deprivation (Fig. 1), whereas O\textsubscript{3} was higher in more affluent areas. The pattern of higher concentrations with increasing deprivation was seen in the North and London, but there was no such trend in the South (eTable 3, http://links.lww.com/EDE/A628). Descriptive statistics for PM\textsubscript{2.5} are shown in eTable 4 (http://links.lww.com/EDE/A628).

Among patients with assigned exposure, the two pollutants most closely related were PM\textsubscript{10} and NO\textsubscript{2} (positive correlation r = 0.86). O\textsubscript{3} was negatively correlated with PM\textsubscript{10}, SO\textsubscript{2}, and NO\textsubscript{2}. Practice was a strong determinant in explaining the amount of variation in pollutant concentrations seen between patients in the cohort (intraclass correlation coefficients in Table 1). For example, 94% of variation for O\textsubscript{3} and 77% of variation for SO\textsubscript{2} were caused by between-practice differences.

Incidence of Cardiovascular Events

A summary of the incident recording of MI, stroke, heart failure, and arrhythmia events is presented in Table 2. For example, for incident MIs, 25,871 patients (3.1%) already had an MI recorded by 1 January 2003 and therefore were excluded from the analyses of incident MI. Among the 810,686 eligible patients, 13,956 (1.7%) had a first MI recorded in 2003 to 2007.

Men were more likely than women to experience an MI, arrhythmia, or heart failure; for stroke, rates were similar. There were strong effects of geography and deprivation, with those in the North and more deprived areas experiencing a greater rate of disease (eg, for MIs, those in the North had an incidence of 2.1% vs. 1.6% in the South; for the most deprived, it was 2.4% vs. 1.4% for the least deprived).

Effects of Air Pollution

The associations between assigned air pollution exposure (for IQR changes in concentrations during 2002) and 5-year incidence of disease (2003–2007) are quantified in a series of HRs in Table 3 (results scaled to 10 µg/m\textsuperscript{3} are given in eTable 5, http://links.lww.com/EDE/A628). The strongest associations with pollution were seen for heart failure.
FIGURE 1. A–D, Assigned exposure in all patients (n = 836,557) by quintiles of age-sex adjusted (modified) Index of Multiple Deprivation (IMD).

TABLE 1. Summary of Assigned Pollutant Levels in 2002 for Study Cohort (n = 836,557)

<table>
<thead>
<tr>
<th>Assigned Pollutant Exposure in 2002</th>
<th>PM_{10}</th>
<th>NO_{2}</th>
<th>SO_{2}</th>
<th>O_{3}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients with linkage, no. (%)</td>
<td>831,788 (99)</td>
<td>831,375 (99)</td>
<td>824,388 (99)</td>
<td>825,598 (99)</td>
</tr>
<tr>
<td>Pollution level (µg/m³)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>19.7 (2.3)</td>
<td>22.5 (7.4)</td>
<td>3.9 (2.1)</td>
<td>51.7 (2.4)</td>
</tr>
<tr>
<td>Range</td>
<td>10.6–29.8</td>
<td>1.7–60.8</td>
<td>0.03–24.2</td>
<td>44.0–66.2</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>3.0</td>
<td>10.7</td>
<td>2.2</td>
<td>3.0</td>
</tr>
<tr>
<td>Region (no. of practices); mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North (n = 81)</td>
<td>19.8 (2.3)</td>
<td>23.4 (6.3)</td>
<td>4.8 (2.1)</td>
<td>50.9 (2.4)</td>
</tr>
<tr>
<td>South (excluding London, n = 96)</td>
<td>19.1 (2.0)</td>
<td>19.4 (6.1)</td>
<td>3.3 (1.9)</td>
<td>52.6 (2.2)</td>
</tr>
<tr>
<td>London (n = 28)</td>
<td>22.5 (1.2)</td>
<td>33.3 (4.5)</td>
<td>3.8 (1.2)</td>
<td>50.2 (0.8)</td>
</tr>
<tr>
<td>Practice in urban area (no. of practices); mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes (n = 173)</td>
<td>20.1 (2.2)</td>
<td>23.6 (7.0)</td>
<td>4.0 (2.0)</td>
<td>51.5 (2.2)</td>
</tr>
<tr>
<td>No (n = 32)</td>
<td>17.7 (1.9)</td>
<td>16.1 (6.2)</td>
<td>3.5 (2.6)</td>
<td>52.3 (3.2)</td>
</tr>
<tr>
<td>Deprivation (modified IMD) quintile; mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (most deprived)</td>
<td>21.0 (2.4)</td>
<td>26.6 (7.1)</td>
<td>4.4 (2.3)</td>
<td>51.1 (2.2)</td>
</tr>
<tr>
<td>2</td>
<td>20.2 (2.5)</td>
<td>23.7 (7.8)</td>
<td>4.4 (2.5)</td>
<td>51.2 (2.2)</td>
</tr>
<tr>
<td>3</td>
<td>19.4 (2.5)</td>
<td>21.4 (7.9)</td>
<td>3.8 (2.2)</td>
<td>51.8 (2.6)</td>
</tr>
<tr>
<td>4</td>
<td>19.3 (2.2)</td>
<td>21.0 (7.0)</td>
<td>3.7 (1.9)</td>
<td>51.7 (2.5)</td>
</tr>
<tr>
<td>5 (least deprived)</td>
<td>19.5 (1.7)</td>
<td>21.9 (6.1)</td>
<td>3.6 (1.5)</td>
<td>52.0 (2.1)</td>
</tr>
</tbody>
</table>

Correlation* with future years exposure

<table>
<thead>
<tr>
<th>Year</th>
<th>PM_{10}</th>
<th>NO_{2}</th>
<th>SO_{2}</th>
<th>O_{3}</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>0.92</td>
<td>0.98</td>
<td>0.81</td>
<td>0.63</td>
</tr>
<tr>
<td>2004</td>
<td>0.91</td>
<td>0.97</td>
<td>0.67</td>
<td>0.55</td>
</tr>
<tr>
<td>2005</td>
<td>0.84</td>
<td>0.97</td>
<td>0.68</td>
<td>0.56</td>
</tr>
<tr>
<td>2006</td>
<td>0.85</td>
<td>0.97</td>
<td>0.61</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Correlation* with other pollutants

<table>
<thead>
<tr>
<th>Pollutant</th>
<th>PM_{10}</th>
<th>NO_{2}</th>
<th>SO_{2}</th>
<th>O_{3}</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO_{2}</td>
<td>0.86</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>SO_{2}</td>
<td>0.52</td>
<td>0.50</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>O_{3}</td>
<td>-0.44</td>
<td>-0.48</td>
<td>-0.45</td>
<td>—</td>
</tr>
</tbody>
</table>

Intraclass correlation (ICC) by practice

<table>
<thead>
<tr>
<th>Year</th>
<th>PM_{10}</th>
<th>NO_{2}</th>
<th>SO_{2}</th>
<th>O_{3}</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>0.87</td>
<td>0.90</td>
<td>0.77</td>
<td>0.94</td>
</tr>
</tbody>
</table>

* Spearman’s rank correlation coefficient.
For example, IQR changes in PM$_{10}$ (3.0 µg/m$^3$) and NO$_2$ (10.7 µg/m$^3$) were both associated with 11% increases in heart failure. These associations were reduced when adjusted for smoking and BMI and further reduced with adjustment for the Index of Multiple Deprivation (HR = 1.06 [95% confidence interval {CI} = 1.01–1.11] for both PM$_{10}$ and NO$_2$). HRs and 95% CIs for PM$_{2.5}$ were similar to those for PM$_{10}$ (eTable 6, http://links.lww.com/EDE/A628).

The evidence for associations of PM$_{10}$ and NO$_2$ with incident MI, stroke, or arrhythmias was weak both before and after adjustment for potential confounders. By contrast, SO$_2$ maintained an association with MIs even after adjustment for the Index of Multiple Deprivation (1.05 [1.02–1.07] for a 2.2-µg/m$^3$ increment in SO$_2$). O$_3$ tended to show negative associations with the outcomes, which lessened with adjustment, although heart failure persisted (6% lower risk with a 3-µg/m$^3$
### TABLE 3. Hazard Ratios* Summarizing the Change in Risk of Incident MI, Stroke, Arrhythmias, and Heart Failure in 2003–2007 Associated with an Interquartile Change in Each Pollutant

<table>
<thead>
<tr>
<th>Pollutant</th>
<th>Baseline Variables</th>
<th>Myocardial Inf. (n = 810,686)</th>
<th>Stroke (n = 819,370)</th>
<th>Arrhythmias (n = 790,751)</th>
<th>Heart Failure (n = 810,188)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>HR* (95% CI)</td>
<td>HR* (95% CI)</td>
<td>HR* (95% CI)</td>
<td>HR* (95% CI)</td>
</tr>
<tr>
<td>PM$_{10}$</td>
<td>Adjusted for age and sex</td>
<td>1.03 (0.99–1.07)</td>
<td>1.02 (0.99–1.05)</td>
<td>1.00 (0.97–1.03)</td>
<td>1.11 (1.06–1.17)</td>
</tr>
<tr>
<td></td>
<td>Further adjusted for smoking, BMI, and comorbidity*</td>
<td>1.01 (0.98–1.05)</td>
<td>1.01 (0.98–1.04)</td>
<td>1.00 (0.97–1.03)</td>
<td>1.09 (1.05–1.14)</td>
</tr>
<tr>
<td></td>
<td>Further adjusted for (modified) IMD</td>
<td>0.98 (0.94–1.01)</td>
<td>0.98 (0.95–1.01)</td>
<td>0.99 (0.96–1.02)</td>
<td>1.06 (1.01–1.11)</td>
</tr>
<tr>
<td>NO$_2$</td>
<td>Adjusted for age and sex</td>
<td>1.04 (0.99–1.09)</td>
<td>1.04 (1.00–1.08)</td>
<td>1.00 (0.96–1.03)</td>
<td>1.11 (1.06–1.17)</td>
</tr>
<tr>
<td></td>
<td>Further adjusted for smoking, BMI, and comorbidity*</td>
<td>1.02 (0.98–1.07)</td>
<td>1.03 (0.99–1.07)</td>
<td>1.00 (0.97–1.03)</td>
<td>1.10 (1.05–1.15)</td>
</tr>
<tr>
<td></td>
<td>Further adjusted for (modified) IMD</td>
<td>0.98 (0.93–1.03)</td>
<td>0.99 (0.95–1.03)</td>
<td>0.99 (0.96–1.02)</td>
<td>1.06 (1.01–1.11)</td>
</tr>
<tr>
<td>SO$_2$</td>
<td>Adjusted for age and sex</td>
<td>1.09 (1.07–1.10)</td>
<td>1.05 (1.04–1.07)</td>
<td>1.03 (1.01–1.04)</td>
<td>1.08 (1.07–1.10)</td>
</tr>
<tr>
<td></td>
<td>Further adjusted for smoking, BMI, and comorbidity*</td>
<td>1.07 (1.04–1.10)</td>
<td>1.04 (1.02–1.07)</td>
<td>1.02 (1.00–1.04)</td>
<td>1.07 (1.03–1.10)</td>
</tr>
<tr>
<td></td>
<td>Further adjusted for (modified) IMD</td>
<td>1.05 (1.02–1.07)</td>
<td>1.02 (1.00–1.05)</td>
<td>1.02 (1.00–1.04)</td>
<td>1.04 (1.01–1.08)</td>
</tr>
<tr>
<td>O$_3$</td>
<td>Adjusted for age and sex</td>
<td>0.93 (0.90–0.96)</td>
<td>0.97 (0.94–1.01)</td>
<td>1.01 (0.98–1.05)</td>
<td>0.91 (0.87–0.95)</td>
</tr>
<tr>
<td></td>
<td>Further adjusted for smoking, BMI, and comorbidity*</td>
<td>0.94 (0.91–0.97)</td>
<td>0.98 (0.95–1.02)</td>
<td>1.01 (0.98–1.05)</td>
<td>0.92 (0.88–0.96)</td>
</tr>
<tr>
<td></td>
<td>Further adjusted for (modified) IMD</td>
<td>0.96 (0.93–1.00)</td>
<td>1.00 (0.97–1.04)</td>
<td>1.02 (0.98–1.05)</td>
<td>0.94 (0.90–0.98)</td>
</tr>
</tbody>
</table>

*Hazard ratios refer to an IQR change in each pollutant (PM$_{10}$ = 3.0 $\mu$g/m$^3$, SO$_2$ = 2.2 $\mu$g/m$^3$, NO$_2$ = 10.7 $\mu$g/m$^3$, O$_3$ = 3.0 $\mu$g/m$^3$).

*Diabetes and hypertension.

**FIGURE 2.** Stratified hazard ratios from fully adjusted model for incidence of heart failure and PM$_{10}$ (A), NO$_2$ (B), SO$_2$ (C), and O$_3$ (D) exposure.
TABLE 4. Within- and Between-Practice Hazard Ratios1 Summarizing the Change in Risk of Incident MI, Stroke, Arrhythmias, and Heart Failure in 2003–2007 From an Interquartile Change in PM$_{10}$

<table>
<thead>
<tr>
<th>Adjustment for Modified IMD</th>
<th>Practice Effect</th>
<th>MI (n = 810,686)</th>
<th>Stroke (n = 819,370)</th>
<th>Arrhythmias (n = 790,751)</th>
<th>Heart Failure (n = 810,188)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>Within</td>
<td>1.08 (1.02–1.15)</td>
<td>1.05 (0.97–1.13)</td>
<td>1.01 (0.97–1.06)</td>
<td>1.09 (1.02–1.16)</td>
</tr>
<tr>
<td>No</td>
<td>Between</td>
<td>0.93 (0.86–1.00)</td>
<td>0.96 (0.89–1.04)</td>
<td>0.98 (0.93–1.04)</td>
<td>1.01 (0.93–1.09)</td>
</tr>
<tr>
<td>Yes</td>
<td>Within</td>
<td>1.01 (0.96–1.07)</td>
<td>1.00 (0.93–1.06)</td>
<td>1.01 (0.96–1.06)</td>
<td>1.02 (0.96–1.09)</td>
</tr>
<tr>
<td>Yes</td>
<td>Between</td>
<td>0.96 (0.90–1.03)</td>
<td>0.98 (0.91–1.06)</td>
<td>0.98 (0.93–1.04)</td>
<td>1.04 (0.97–1.12)</td>
</tr>
</tbody>
</table>

1All models adjusted for age, sex, smoking, BMI, diabetes and hypertension. Hazard ratios refer to an IQR change in PM$_{10}$ of 3.0 µg/m$^3$.

increase). Results for the broader groups of CHD and cerebrovascular disease were similar to the respective subgroups of MI and stroke (eTable 7, http://links.lww.com/EDE/A628).

Fitting a time-varying covariate for pollution in the model based on previous years’ exposure through the study made no appreciable difference to our results (data not shown). We also investigated the robustness of effect sizes to adjustment for other pollutants (eTable 8, http://links.lww.com/EDE/A628). The full model effects of a 6% increase in heart failure for an IQR change in PM$_{10}$ and NO$_2$ reduced to 5% when SO$_2$ was added to the model and to 3 to 4% when O$_3$ was added. The negative associations with O$_3$ moved toward unity with adjustment for all pollutants.

Effect Modification

Figure 2 shows the results from fitting the fully adjusted model for heart failure for all pollutants stratified by sex, age, BMI, smoking, Index of Multiple Deprivation, and region. For PM$_{10}$, NO$_2$, and SO$_2$, associations tended to increase as the level of deprivation decreased. The pattern among other stratified variables was less clear, although patients with BMI > 30 and ex-smokers had smaller associations for all pollutants. There was some variation in the effect size of the associations by geography; larger associations were seen in the South than in the North. Within London practices, associations were greater (except for NO$_2$), although CIs were wide.

Clustering Effect of Practice

We investigated the influence of practice by partitioning the HR into between- and within-practice estimates. Table 4 shows the results of this for PM$_{10}$ in models with and without adjustment for Index of Multiple Deprivation. In the model not adjusting for the Index of Multiple Deprivation, positive associations are seen within practice (ie, 8% increase in MIs per IQR increase in PM$_{10}$, which contrasts with negative associations between practice (7% decrease in MIs per IQR increase in PM$_{10}$). The adjustment for the Index of Multiple Deprivation brings these divergent associations much closer together, removing most of the within-practice effects. A similar pattern was noted with the other pollutants (data not shown).

DISCUSSION

In this large, population-based cohort study, we found inconsistent evidence to support the hypothesis that long-term exposure to particulate air pollution was associated with increased incidence of CVD. Associations between PM and incident CHD, MI, stroke, and arrhythmias were close to unity but increased for heart failure. Increased incidence of heart failure was also positively associated with NO$_2$ and SO$_2$.

Our study is the first to investigate the long-term effects of air pollution on the development of CVD in a large UK population incorporating individual risk factors. Previously, Elliott et al28 reported positive associations between black smoke and SO$_2$ and subsequent mortality in Great Britain using small-area population and socioeconomic status statistics, but lacking data on individual risk factors. More recently, a large representative study of the UK population reported on long-term effects in relation to the prevalence of CHD and found small associations with PM$_{10}$.

Air Pollution Exposure Assignment

In our study, pollution estimates for 1 × 1 km grids across England (mapped to residents’ postcodes) were estimated using pollution emission inventories and meteorological information combined with air dispersion models. Dispersion models, together with land-use-regression models that predict pollution concentrations at a given site based on surrounding land use and traffic characteristics, provide improvements on methods that assign exposure using measurements from the nearest monitor.21 It has been suggested that dispersion models are more sophisticated and reliable than land-use-regression models in intrarural settings.30 The key advantage of dispersion models over other approaches is that they provide a better representation of the process under study.30 However, they are more expensive to implement because of their substantial data requirements.30 A comparison of the performance of land-use-regression and dispersion models in predicting annual NO$_2$ concentrations in the Netherlands concluded there was moderate agreement between the methods.31 Furthermore, the performance of our models was comparable to the land-use-regression data available for the United Kingdom.32,33
In our study, the validity of the modeled exposure data varied among the pollutants and from year to year \((R^2\) statistics for each pollutant for years 2002–2007 are given in the eAppendix, http://links.lww.com/EDE/A628), both in relation to the national network sites and the smaller number of verification sites. Although the \(R^2\) values based on verification sites provide the strongest indication of the performance of the model, the manner in which the national network data are used in the calibration process does not preclude the national network sites from providing some indication of model performance. The difficulty in modeling PM is well established\(^{33}\) because of the complexity of the PM mixture for example. The removal of sulfur from petrol and diesel as a result of European Union legislation and the move away from coal to gas and electricity for domestic and industrial use have caused \(SO_2\) levels to fall substantially in the last 50 years, making them harder to model.\(^{32}\) Our validation statistics for \(PM_{10}\) and \(SO_2\) were broadly similar to UK estimates derived using land-use-regression models for \(PM_{10}\) concentrations in 2001\(^{33}\) and \(SO_2\) concentrations in 1991.\(^{32}\) Nonetheless, we believe the poorer model performance for \(PM_{10}\) and \(SO_2\) relative to \(NO_2\) should be a consideration in the overall assessment of our results.

Despite the year-to-year variability in the validation \(R^2\) statistics, our sensitivity analysis demonstrated that the estimated HRs were robust to the choice of exposure year. Furthermore, a time-varying-exposure survival model gave results comparable to those derived from 2002 data alone (data not shown). We are therefore confident in the choice of 2002 as the exposure period used for our study, despite the lower \(R^2\) statistics compared with other years.

Misclassification of grid square pollutant concentration by the model is only one potential source of error in our exposure assignment. Another is misclassification of personal exposure arising from the use of exposures based solely on residential address, rather than monitored personal exposure (from other sources such as indoor or workplace). We acknowledge these weaknesses in our study, but note that it is likely that this type of measurement error would bias effect estimates toward the null.\(^{34}\) However, we were able to demonstrate robust associations between all pollutants and incidence of heart failure.

### Studies of PM and Cardiovascular Events

Evidence linking PM to the incidence of both fatal and nonfatal CVD in large population cohorts has been limited to three studies of women in the United States—the Women’s Health Initiative,\(^{20}\) the Nurses’ Health Study,\(^{15}\) and the California Teachers Study.\(^{17}\) All three studies identified events from a combination of questionnaires, medical records, and death registration. Miller et al\(^{20}\) reported an increased HR for \(PM_{10}\) and first cardiovascular event (MI, coronary revascularization, stroke, CHD, and cerebrovascular death) of 1.04 (95% CI = 0.95–1.12) per 10 \(\mu g/m^3\). For MI, the HR for \(PM_{2.5}\) was 1.06 (0.85–1.34), and for stroke, 1.28 (1.02–1.61) per 10 \(\mu g/m^3\) increment in \(PM_{2.5}\). Puett et al\(^{15}\) reported increased HRs for \(PM_{2.5}\) and \(PM_{2.5–10}\) for incident CHD events but with lower CIs below 1. Lipsett et al\(^{35}\) reported little evidence for an association between either \(PM_{10}\) or \(PM_{2.5}\) and MI incidence (HR 0.98 [0.91–1.06] and 0.98 [0.83–1.16], respectively) but a stronger association for both \(PM_{10}\) and \(PM_{2.5}\) and stroke (1.06 [1.00–1.13] and 1.14 [0.99–1.32], respectively) each per 10 \(\mu g/m^3\) increment. In our study of more than 800,000 English adults, we found HRs for \(PM_{10}\) and \(PM_{2.5}\) with incident CHD, MI, and stroke that were close to unity, consistent with the rather weak evidence for MI, but inconsistent with the evidence for the stronger associations for stroke reported in the studies referenced above.\(^{15,17,20}\) Of the many methodological differences with our work, we note that the modeled levels of particulate exposure were lower in the United Kingdom, and that these US studies were based on women alone. Although it has been suggested that the associations between PM and CHD mortality are stronger in women than in men,\(^{35}\) we found no evidence of this in our study.

Our study found associations between the incidence of heart failure events and residential levels of \(PM_{10}\) and \(PM_{2.5}\) exposure. This finding extends previous evidence linking long-term exposure to PM and mortality from heart failure.\(^{19,36}\) Pope et al\(^{15}\) analyzed data from the American Cancer Society cohort and reported an increased HR for mortality from heart failure (including dysrhythmias and cardiac arrest) of 1.13 (95% CI = 1.05–1.21) per 10 \(\mu g/m^3\) increment in \(PM_{2.5}\). Beelen et al\(^{16}\) assessed the association between \(PM_{2.5}\) and heart failure mortality in the Netherlands Cohort Study on Diet and Cancer and reported an HR larger than ours (2.69 [95% CI = 1.37–5.27] per 10 \(\mu g/m^3\) increment). There is a more substantial body of evidence linking short-term exposure to particulate air pollution and heart failure mortality and morbidity.\(^{2}\)

### Results for Gases

Few studies have reported results for gaseous pollutants and cardiovascular outcomes. Miller et al\(^{20}\) observed positive associations with \(NO_2\), \(O_3\), and \(SO_2\), with CIs that included 1.0. Rosenlund et al\(^{17}\) reported positive associations between \(NO_2\) and \(SO_2\) and incident MI, again with CIs that included 1.0, in a case-control study of first-time MI cases and population controls age 45 to 70 years in Stockholm county. Gan et al\(^{16}\) and Rosenlund et al\(^{13}\) focused on traffic sources in their studies, observing positive associations between \(NO_2\) and CHD events with lower confidence limits close to 1.0. These findings are broadly consistent with the general lack of evidence for an association with \(NO_2\) and \(O_3\) in our study. Our finding of consistent associations between \(SO_2\) and increased incidence of a range of CVDs was in line with the results from the only other studies from the United Kingdom that assessed long-term exposure to air pollution and disease.\(^{28,29}\) \(SO_2\) has also been associated with increased numbers of deaths from
all disease-related causes, cardiopulmonary causes, and lung cancer. However, given the lack of biological plausibility for health effects of low concentrations of SO₂, the pollutant may simply be an indicator for combustion sources.

In our study, we observed either negative or null associations between O₃ and CVD events. Because O₃ is a regional pollutant, variability in O₃ concentrations is mainly between regions, with little within-practice variation. Ozone is highest in the South where disease incidence is lowest, and lowest in the North where disease incidence is highest, and therefore, regional differences are likely to explain the observed relationships. In addition, O₃ has a pronounced seasonal pattern, which our annual estimates do not reflect. Stronger associations have been observed in studies using summer-only concentrations. However, these data were unavailable to us, which is a limitation of our study.

Adjustment for Socioeconomic Status

We presented estimates adjusted for our marker of socioeconomic status (modified Index of Multiple Deprivation), which had the effect of moving all associations toward the null. We think the adjustment is potentially important because it controls for many unmeasured confounders, such as diet and nutrition. However, because increasing deprivation was associated with higher levels of pollution (except ozone) in our study, there is a chance that this adjustment is removing effects of air pollution. The finer spatial resolution of the Index of Multiple Deprivation variable in some urban areas may represent fine-scale variability in the pollution levels not captured in the 1 × 1 km resolution of the pollutant estimates. Other studies have adjusted for socioeconomic status using a variety of indicators. Some used household income and reported little impact on pollutant associations. Others have used small-area Census-based measures of deprivation and observed increased estimates of the effects of air pollution on coronary events and mortality. Concluded that this was likely representing an adjustment for smoking (for which they had no data) but also noted that in Rome the most affluent were living in the most polluted areas, a finding also observed in the American Cancer Society study. This was not the case in our study, in which PM₁₀ concentrations generally increased with deprivation.

Effect Modification

We did not replicate other authors’ findings of effect modification with factors such as socioeconomic status and BMI; larger associations have been reported in the less educated and more obese. Instead, we tended to show the reverse, with greater overall effects seen in those who were least deprived and of normal weight. We attribute this to the fact that the region that consistently showed associations (the South) had a larger proportion of patients who were the least deprived and of normal weight (eg, Index of Multiple Deprivation least deprived quintile: 34% of patients in the South vs. 14% in the North). However, we also observed that the overall relationship between deprivation and pollution (the most deprived patients tended to reside in areas with higher levels of pollution) was not apparent in the South. A Swedish study of pollution and MI found larger effects in the better educated; the authors speculated that this may be due to the fact that in Stockholm more socioeconomically privileged people tended to live nearer the city centre. We acknowledge that the differing nature of the relationship between socioeconomic status and pollution across England may account for the geographical variations we observed.

Data Sources

Our study was strengthened by using three independent sources of data to identify incident events. To ensure data quality, the Clinical Practice Research Datalink protocol specifies internal checks on data continuity and completeness and provides recording guidelines to all practices that contribute data to it. The introduction in 2004 of the Quality and Outcomes Framework (a “pay-for-performance” contract for primary care) further improved the recording of CVD on primary care databases. In the United Kingdom, there is complete coverage for death registrations, admissions to NHS hospitals, and NHS-funded admissions to private hospitals. Patients from practices consenting to data linkage (approximately 40% of current practices) are representative of all Clinical Practice Research Datalink patients. Our analysis of these data sources suggests that studies using only hospitalization or mortality records may miss some events because of coding differences and inconsistencies among the systems (data not shown). The misclassification of events (eg, when an initial diagnosis by the general practitioner is not confirmed at admission to hospital) cannot be ruled out, however.

Conclusion

Our study found evidence of an association between long-term exposure to particulate pollution and NO₂ with the incidence of heart failure events but not with MI, arrhythmias, or stroke. These results, from one of the few studies to be based on a whole population sample, differ from those reported from the United States and mainland Europe. Further research is required to understand the reasons for these differences. Additional verification of SO₂ and PM₁₅ data is required to further understand the role of these pollutants in the onset of CVD.

Acknowledgments

The Clinical Practice Research Datalink (CPRD) is owned by the Secretary of State of the UK Department of Health and operates within the Medicine and Healthcare Products Regulatory Agency (MHRA). CPRD has received funding from the MHRA, Wellcome Trust, Medical Research Council, NIHR Health Technology Assessment Programme, Innovative Medicines Initiative, UK Department of Health, Technology Strategy Board, Seventh Framework Programme EU, various universities, contract research organizations, and pharmaceutical companies.
REFERENCES


