Residential Air Pollution and Associations with Wheeze and Shortness of Breath in Adults: A Combined Analysis of Cross-Sectional Data from Two Large European Cohorts

Dany Doiron,1,2,3 Kees de Hoogh,2,3 Nicole Probst-Hensch,2,3 Stéphane Mboutchou,1 Marloes Eeftens,4 Yutong Cai,5 Christian Schindler,4,5 Isabel Fortier,1 Susan Hodgson,5 Amadou Gaye,6 Ronald Stolk,7 and Anna Hansell8

1Research Institute of the McGill University Health Centre, Montreal, Quebec, Canada
2Swiss Tropical and Public Health Institute, Basel, Switzerland
3University of Basel, Basel, Switzerland
4Institute for Risk Assessment Sciences, Utrecht University, Utrecht, Netherlands
5MRC-PHE Centre for Environment and Health, Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK
6National Human Genome Research Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland, USA
7University Medical Center Groningen, Groningen, Netherlands

BACKGROUND: Research examining associations between air pollution exposure and respiratory symptoms in adults has generally been inconclusive. This may be related in part to sample size issues, which also preclude analysis in potentially vulnerable subgroups.

OBJECTIVES: We estimated associations between air pollution exposures and the prevalence of wheeze and shortness of breath using harmonized baseline data from two very large European cohorts, Lifelines (2006–2013) and UK Biobank (2006–2010). Our aim was also to determine whether the relationship between air pollution and respiratory symptom prevalence differed between individuals with different characteristics.

METHODS: Cross-sectional analyses explored associations between prevalence of self-reported wheeze and shortness of breath and annual mean particulate matter with aerodynamic diameter <2.5 μm, 2.5–10 μm, and <10 μm (PM2.5, PM10, and PM2.5, respectively) and nitrogen dioxide (NO2) concentrations at place of residence using logistic regression. Subgroup analyses and tests for interaction were performed for age, sex, smoking status, household income, obesity status, and asthma status.

RESULTS: All PM exposures were associated with respiratory symptoms based on single-pollutant models, with the largest associations seen for PM2.5 with prevalence of wheezing [odds ratio (OR) = 1.16 per 5 μg/m3 [95% confidence interval (CI): 1.11, 1.21]] and shortness of breath [OR = 1.61 per 5 μg/m3 (95% CI: 1.45, 1.78)]. The association between shortness of breath and a 5-μg/m3 increment in PM2.5 was significantly higher for individuals from lower-income households [OR = 1.31 (95% CI: 1.11, 1.55); p-interaction = 0.005], whereas the association between PM2.5 and wheeze was limited to lower-income participants [OR = 1.30 (95% CI: 1.22, 1.38) vs. OR = 1.02; (95% CI: 0.96, 1.08); p-interaction <0.001]. Exposure to NO2 also showed positive associations with wheeze and shortness of breath.

CONCLUSION: Exposure to PM and NO2 air pollution was associated with the prevalence of wheeze and shortness of breath in this large study, with stronger associations between PM2.5 and both outcomes among lower- versus higher-income participants.

Introduction

In 2010, ambient particulate matter (PM) air pollution was ranked as the ninth highest attributable burden risk factor and was estimated to cause 3.1 million premature deaths and 3.1% of disability adjusted life years (DALYs) worldwide (Lim et al. 2012). Chronic obstructive pulmonary disease (COPD), acute lower respiratory disease (ALRD), and lung cancer accounted for 28% of deaths attributable to ambient air pollution in 2012 (WHO 2016). Ambient air pollution is also thought to increase the risk of asthma exacerbation and asthma onset in children and adults (Guarnieri and Balmes 2014).

Respiratory symptoms such as wheezing and shortness of breath are indicators of airway inflammation associated with chronic respiratory diseases such as COPD, emphysema, and asthma, among other conditions (Abramson et al. 2002). Wheezing is caused by narrowed or compressed lower airways, which leads to turbulent airflow (Loudon and Murphy 1984). Shortness of breath is broadly defined as a subjective feeling of breathing discomfort. It can be characterized by a sense of respiratory effort/work, chest tightness, “air hunger/unsatisfied inspiration, or a combination of these symptoms (Parshall et al. 2012). Wheezing usually relates to respiratory dysfunction, whereas breathlessness can be caused by cardiac as well as by respiratory dysfunction (Bozkurt and Mann 2003). Smoking and environmental tobacco smoke are important risk factors for COPD and for asthma development and symptoms (Janson et al. 2001; Leuenberger et al. 1994). Like cigarette smoke, ambient air pollution is a source of particulate matter exposure, but research on the relationship between air pollution and wheeze and shortness of breath prevalence in adults has largely been inconsistent (Brunkereo and Holgate 2002) despite evidence for air pollution impacts on respiratory mortality and asthma exacerbation (Fischer et al. 2015; Guarnieri and Balmes 2014).

Ambient air pollution was associated with incident asthma in a follow-up study of never smokers in the Swiss Study on Air Pollution and Lung Disease in Adults (SAPALDIA) cohort (41 cases) (Künzli et al. 2009), and authors of a combined longitudinal analysis of six European cohorts (including SAPALDIA) reported weak positive associations between air pollution and the incidence of adult asthma (1,257 cases) (Jacquemin et al. 2015). Data collected from >8,500 SAPALDIA participants in 1991 and 2002 suggested that living near a major street or highway was positively associated with attacks of breathlessness in the previous 12 months (Bayer-Oglesby et al. 2006). Residential ambient...
air pollutant concentrations were also associated with the prevalence of breathlessness, but not wheeze, in a cross-sectional study of 9,651 adults in the SAPALDIA cohort (Zemp et al. 1999). However, outdoor air pollution was associated with incident wheeze (1,023 cases) and incident asthma (254 cases) in a longitudinal analysis of a cohort of adult women in the United States (Young et al. 2014), and a study of 2,628 U.S. male veterans reported an association between residential proximity to a major roadway and persistent wheeze during the past year (Garshick et al. 2003).

Lastly, a cross-sectional survey in Netherlands reported that 673 adults living along busy streets were more likely to report shortness of breath during walking, but not wheeze, than 812 adults living along streets in the same neighborhoods that had little traffic (Oosteree et al. 1996). Many of these studies had limited numbers of participants (and cases), resulting in imprecise estimates of associations between air pollution exposures and respiratory health outcomes and in an inability to explore factors that might modify associations.

PM and nitrogen dioxide (NO₂) are commonly used as markers of ambient air pollutant exposures for epidemiologic studies. In the context of the BioSHaRE (Biobank Standardisation and Harmonisation for Research Excellence in the European Union) program, we harmonized and combined data from two of Europe’s largest population health studies, the Lifelines Cohort Study and Biobank (Stolk et al. 2008) and UK Biobank (UK Biobank 2007), to explore associations of residential PM and NO₂ exposures with the prevalence of wheeze and shortness of breath in cohorts as a whole and in potentially vulnerable population subgroups. To our knowledge, this is the largest study of associations between respiratory symptoms and air pollution to date.

**Methods**

**Study Populations**

Both the Lifelines and UK Biobank cohorts sampled their participants from the general population using general practice registries. Baseline questionnaire and physical measures data were collected from 152,180 adult Lifelines participants (18–93 y of age) during 2006–2013 (Stolk et al. 2008) and from 502,655 UK Biobank participants (40–69 y of age) during 2006–2010 (UK Biobank 2007). The Lifelines cohort was designed to include three generations of study participants. An initial set of participants was identified from general practice patients residing in the three northernmost provinces of Netherlands (Friesland, Groningen, and Drenthe) which are largely rural regions (Klijns et al. 2015; Scholtens et al. 2015). The initial group of eligible participants comprised adults who were 25–50 y of age who did not have a severe psychiatric or physical illness or a limited life expectancy and who were able to complete the Dutch-language study questionnaire. Participants were asked to identify family members who might also want to participate in the study, and adult residents of the study provinces also could volunteer directly to participate in the study. Individuals were identified and invited to join the UK Biobank cohort via population-based registries such as those held by the National Health Service. Invitations were stratified according to key sociodemographic factors (e.g., age, sex, and postcode areas as a measure of social deprivation), with over-sampling of particular groups to enhance generalizability and to account for the impact of participation rates (UK Biobank 2007). UK Biobank study participants lived within approximately 25 miles of one of 22 assessment centers across Scotland, England, and Wales, in mostly urban areas (Allen et al. 2012). Full study sampling methods and participant selection criteria have been defined elsewhere (Allen et al. 2012; Scholtens et al. 2015; Stolk et al. 2008; UK Biobank 2007). Participants in both studies provided written informed consent at recruitment for broad use of their data by local and international investigators. We submitted research protocols and data access applications to the Lifelines and UK Biobank ethics and scientific review boards and obtained all necessary approvals.

**Standardized Air Pollution Data**

We used annual average concentrations of NO₂ and of particulate matter with aerodynamic diameter <10 μm (PM₁₀), fine particles with diameter <2.5 μm (PM₂.₅), and coarse particles with diameter between 2.5 μm and 10 μm (PM₁₀coarse) at place of residence for the periods 2009–2010 (Lifelines) and 2010–2011 (UK Biobank). These air pollution estimates were generated in the context of the European Study of Cohorts for Air Pollution Effects (ESCAPE) using a standardized land-use regression (LUR) modeling approach. ESCAPE LUR models developed for Netherlands/Belgium and for southeast England (i.e., London and Oxford) were employed for the Lifelines and UK Biobank study areas, respectively. Geographic information systems (GIS)-derived predictor variables were used to model spatial variation of measured annual air pollutant concentrations in each area, and model performance was evaluated using leave-one-out cross-validation (see Tables S1 and S2). Modeled PM₂.₅, PM₁₀, and NO₂ performed relatively well in the Netherlands/Belgium area (cross-validation $R^2 = 61\%$, 60%, and 80%, respectively) and in the southeast England area (cross-validation $R^2 = 82\%$, 90%, and 89%, respectively). PM₁₀coarse cross-validation $R^2$ results were lowest in both areas (38% in the Netherlands/Belgium area; 57% in the southeast England area). Details of the ESCAPE LUR model development and validation methodology have been described elsewhere (Beelen et al. 2013; Eeftens et al. 2012).

For our study, home addresses of Lifelines and UK Biobank participants at the baseline assessment were geocoded and linked to LUR ESCAPE air pollution estimates. Because a significant proportion of UK Biobank participants resided in areas outside the initial southeast England ESCAPE study area, we tested the transferability of the LUR models using historical monitoring data from the United Kingdom’s Automatic Urban and Rural Network (AURN) (Gulliver and de Hoogh 2015). The southeast England ESCAPE LUR model was applied to all UK AURN site locations, and associations between the LUR-modeled and AURN-measured NO₂ and PM₁₀ concentrations were analyzed by distance bands radiating from the edge of the ESCAPE study area.

**Data Harmonization**

To ensure data compatibility in our project, all respiratory symptoms and covariates used in statistical analyses went through a structured data harmonization process (Doiron et al. 2013; Fortier et al. 2016). As a first step, common-format respiratory outcome and confounding variables required for coanalyses were identified, and the possibility for each cohort to generate these variables was determined using information extracted from questionnaires, data dictionaries (i.e., codebooks), and standard operating procedures. Once "target" variables were defined and the harmonization potential was determined, processing algorithms written in the JavaScript programming language were implemented in Opal data harmonization software (Doiron et al. 2017) to map data collected from each cohort to a common (i.e., harmonized) format. Lastly, to validate the harmonized data sets, we conducted consistency and logic checks to verify counts and distributions of missing data, value ranges, and recoded variables.

Respiratory symptoms were collected through self-administered surveys at the baseline assessments for Lifelines and UK Biobank using different questionnaire assessment items (see Table S3). Both
studies collected data on prevalence of wheeze and shortness-of-breath symptoms. Lifelines collected nonperiod-specific prevalence of wheeze (“Have you had wheezing or whistling in your chest at any time?”), wheezing without a cold (“Have you had this wheezing or whistling when you did not have a cold?”), and wheezing combined with shortness of breath (“Have you been at all breathless when the wheezing noise was present?”). UK Biobank collected information on wheezing in the last year (“In the last year, have you ever had wheeze or whistling in the chest?”). Lifelines participants reported shortness of breath at rest, whereas the UK Biobank questionnaire asked about shortness of breath while walking on level ground. Neither cohort included a recall period for shortness of breath, and only ~35% of UK Biobank participants had data for this question because it was not added to the baseline survey until 2009. To account for heterogeneity of study-specific assessment items, the harmonized wheeze and shortness-of-breath symptom variables were respectively defined as follows: “presence of wheezing symptoms in the past year or more” and “shortness of breath at rest or when walking on level ground” (see Table S4).

We identified sociodemographic and behavioral risk factors to include as potential confounders and for exploring subgroups of vulnerability through literature review and based on data collected from each cohort (see Table S3). Common-format baseline assessment variables for age, sex, body mass index (BMI) (derived from measured height and weight), household income, highest level of education attained, current passive smoking exposure, and smoking status were defined and harmonized across cohorts (Table S4). For example, we performed the following procedures to derive a common-format variable for disposable household income. First, we derived a net annual household income variable for each cohort. For Lifelines, this variable was based on responses to a question about average monthly household income after taxes. For UK Biobank participants, we adjusted self-reported information on gross annual household income to reflect net income by subtracting estimates of participants’ annual tax payments, which were obtained from the UK Government Web Archive (Directgov 2009). Next, we used data from the Organisation for Economic Co-operation and Development (OECD) (http://stats.oecd.org/) to derive a dichotomous income variable defined as “less than or equal to” or “greater than” the country-specific net mean household income for the United Kingdom (20,585 GBP) and Netherlands (25,600 EUR) in 2010.

**Statistical Analysis**

Harmonized data were analyzed using the DataSHIELD federated analysis method (Gaye et al. 2014; Wolfson et al. 2010). DataSHIELD involves setting up secure servers hosting harmonized data in geographically dispersed research centers and allowing a central analysis computer to conduct remote statistical analyses of individual-level data without the need to physically pool the data. DataSHIELD relies on three main components, all of which are open-source and freely available: Opal data management and harmonization software (Doiron et al. 2017), the R statistical programming environment (version 3.3.2; R Development Core Team), and DataSHIELD-specific R packages. For this project, harmonized data sets were hosted on separate servers at the University Medical Center Groningen in Netherlands (Lifelines data) and the Research Institute of the McGill University Health Centre in Canada (UK Biobank data). For an in-depth description of DataSHIELD implementation in the BioSHaRE project, see Gaye et al. (2014).

We used separate logistic regression models to estimate associations between annual average exposures to PM$_{2.5}$, PM$_{10}$, and NO$_2$ at each participant’s residence and self-reported wheezing (in the last year for UK Biobank, at any time for Lifelines; yes/no) and shortness of breath (when walking on level ground for UK Biobank, at rest for Lifelines, with no time period specified for either study; yes/no). For each respiratory symptom, we conducted pooled and study-specific logistic regression models in sequential stages, gradually adjusting for potential confounders to better examine their impact on effect estimates. To facilitate comparisons, these models were restricted to participants with data for all covariates in the fully adjusted model. A first minimally adjusted model included age as a continuous variable and sex (male/female) as confounders. In a second model, we further adjusted for BMI (kilograms per meter squared) [as normal (BMI <25), overweight (BMI = 25 to <30), or obese (BMI ≥30), where BMI was calculated from objectively assessed height and weight], highest level of education attained (less than or equal to secondary education/greater than postsecondary education), and annual net household income (less than or equal to/ greater than the country-specific mean for disposable household income in 2010). The fully adjusted model added smoking status (nonsmoker/past smoker/current smoker) and passive smoking (not exposed at home or work/exposed) to the second model. All pooled models also included a study indicator variable. Using the fully adjusted model, we conducted sensitivity analyses by restricting analyses to study participants living at the same (baseline) address, for ≥5 years and ≥10 years. We then conducted two-pollutant models by adding NO$_2$ as a confounder to the fully adjusted PM regression models and adding PM$_{2.5}$ as a confounder to the NO$_2$ model. We also explored the prevalence of wheeze and shortness of breath according to smoking status and passive smoking exposure to investigate the accuracy of symptom reporting, assuming that individuals exposed to tobacco smoke would have higher symptom prevalence.

Using the third fully adjusted model, we carried out subgroup analyses and tests for interaction to explore whether associations with residential air pollution exposure differed between individuals with different characteristics. Analyses stratified according to age (<65 vs. ≥65 y), sex, smoking status (never smoked vs. current/former smokers), household income (<country-specific mean disposable household income vs. >country-specific mean disposable household income), obesity status [obese (BMI ≥30 kg/m$^2$) vs. nonobese], and self-reported asthma status (“ever had asthma” vs. “never had asthma”) were performed. We tested potential effect modification by including interaction terms in study-specific and pooled analyses. Given potential confounding due to differences in the effects of subgroup variables on respiratory symptoms across studies, we added a second interaction term between the study indicator variable and the respective subgroup variable to the pooled model. Based on the existing literature, we hypothesized increased susceptibility to air pollution among the elderly, smokers, lower-income individuals, obese individuals, and individuals with asthma. Results were considered statistically significant at a level of $p < 0.05$ throughout.

**Results**

When compared with historical monitoring data, the southeast England ESCAPE NO$_2$ model predicted concentrations reasonably well ($R^2 = 0.67$) across the United Kingdom going back to 2006. For PM$_{10}$, the LUR model predicted concentrations moderately well up to 400 km away from the initial ESCAPE study area ($R^2 = 0.53$). All PM analyses were therefore restricted to UK Biobank residential addresses within 400 km of the southeast England area.

Figure 1 shows flow diagrams outlining study populations, exclusions, and missing data. By the time of the present study, the home addresses of 82,959 (54.5%) of the 152,180 participants...
had not yet been geocoded by Lifelines data managers (geocoding was performed in the order of participant recruitment). For UK Biobank, 7,393 (1.5%) of the 502,655 residential addresses could not be geocoded. An additional 33,934 residential addresses were excluded from all PM measures because they were farther than 400 km away from the ESCAPE monitoring area in southeast England. Residential NO2 and PM exposure estimates were therefore available for 495,262 and 461,328 UK Biobank participants, respectively. The household income variable was missing for 17% of Lifelines participants and for 15% of UK Biobank participants, and 16% of UK Biobank participants had no data for the passive smoking exposure variable. Finally, only 35% of UK Biobank participants had data for shortness of breath because this item was added to the baseline survey late in the recruitment phase (i.e., as of 2009).

Population characteristics and respiratory symptom prevalence for each cohort are presented in Table 1. Women comprised 57 and 53% of the Lifelines and UK Biobank participants, respectively. On average, UK Biobank participants were older, had higher BMI, and had lower educational attainment than Lifelines participants. The mean household income levels were comparable across study populations. There was a higher proportion of current smokers and of individuals exposed to secondhand smoke in Lifelines than in UK Biobank. The prevalence of self-reported respiratory symptoms was approximately the same across cohorts, with 1 in 5 participants reporting wheeze and 1 in 10 reporting shortness of breath. Pooled data showed the prevalence of wheeze was higher in former smokers (20.93%) and current smokers (27.08%) than in never smokers (16.43%), and it was higher in those exposed versus not exposed to passive smoke (23.84% vs. 17.22%) (see Table S5). Shortness of breath was also more common in former and current smokers (11.07% and 11.52%, respectively) than in never smokers (8.85%) and in those exposed versus not exposed to passive smoke (13.45% vs. 8.8%). These patterns were also evident for the individual cohorts (see Table S5).

Table 2 presents descriptive statistics for air pollutants. On average, Lifelines participants were exposed to higher PM and lower NO2 concentrations than UK Biobank participants. For all air pollutants, the UK Biobank study area had a wider range. In both studies, annual mean concentrations of PM10 were highly correlated with PM2.5 (Lifelines $r=0.87$; UK Biobank $r=0.81$) and PM1.5 (Lifelines $r=0.78$; UK Biobank $r=0.53$). Annual mean NO2 concentrations were highly correlated with PM2.5 in UK Biobank ($r=0.87$) but not in Lifelines ($r=0.42$) and with PMcoarse in Lifelines ($r=0.77$) but not in UK Biobank ($r=0.19$) (Table 2).

The addition of confounders between pooled logistic regression models 1 through 3 gradually reduced the size of associations between air pollutants and respiratory symptoms (see Table S6) for both addition of socioeconomic status (SES) and BMI variables (model 1 to 2) and for smoking variables (model 2 to 3). The relative degree of attenuation was greater for wheeze than for shortness of breath, although PM2.5 and shortness of breath showed the largest absolute change in the odds ratio (OR) from the minimally adjusted OR of 2.10 (model 1) to 1.69 (model 2) to 1.61 (model 3). Fully adjusted pooled and cohort-specific results are shown in Table 3. In the pooled analyses, exposure to PM2.5 was positively associated with wheeze [OR = 1.16 per 5 µg/m³ (95% CI: 1.11, 1.21)] and shortness of breath [OR = 1.61 per 5 µg/m³ (95% CI: 1.45, 1.78)]. PMcoarse exposure showed a positive nonsignificant association with wheeze prevalence [OR =
Table 1. Population characteristics and respiratory symptom prevalence.

| Characteristic                      | Lifelines n = 52,064 | Wheeze and NO2 analyses; n = 350,639 | Shortness of breath and NO2 analyses; n = 121,550 |
|------------------------------------|----------------------|--------------------------------------|--------------------------------------------------|
| Male, % (n)                        | 43.2 (22,487)        | 47.3 (165,717)                       | 47.3 (57,486)                                    |
| Female, % (n)                      | 56.8 (29,577)        | 52.7 (184,922)                       | 52.7 (64,064)                                    |
| Age, mean ± SD                     | 42.6 ± 11.5          | 56.2 ± 8.1                           | 56.5 ± 8.1                                       |
| BMI, % (n)                         |                       |                                     |                                                  |
| <65 y                               | 96.5 (50,233)        | 82.5 (289,278)                       | 81.2 (98,689)                                    |
| ≥65 y                               | 3.5 (1,831)          | 17.5 (61,361)                        | 18.8 (22,861)                                    |
| Smoking status, % (n)              |                       |                                     |                                                  |
| Current smoker                     | 23.7 (12,345)        | 3 (10,453)                           | 3 (6,670)                                        |
| Former smoker                      | 30.4 (15,813)        | 38.1 (133,703)                       | 38 (46,191)                                      |
| Never smoker                       | 45.9 (23,960)        | 58.9 (206,483)                       | 59 (71,689)                                      |
| Passive smoking exposure, % (n)    |                       |                                     |                                                  |
| Not exposed to secondhand smoke at home or at work | 73.5 (38,264) | 77.4 (272,604) | 77.9 (94,708) |
| Exposed to secondhand smoke at home or at work | 26.5 (13,800) | 22.6 (78,035) | 22.1 (26,842) |
| Wheeze, % (n)                      |                       |                                     |                                                  |
| Has not had wheeze symptoms        | 79.7 (41,466)        | 81.6 (286,276)                       | —                                                 |
| Has had wheeze symptoms            | 20.4 (10,596)        | 18.4 (64,363)                        | —                                                 |
| Shortness of breath, % (n)         |                       |                                     |                                                  |
| Has not had shortness of symptoms  | 90.2 (46,859)        | —                                    | 90.2 (109,584)                                   |
| Has had shortness of symptoms      | 10 (5,205)           | —                                    | 9.8 (11,966)                                     |
| Asthma, % (n)                      |                       |                                     |                                                  |
| Has had asthma                     | 8.1 (4,203)          | 11.6 (40,680)                        | 11.5 (13,967)                                    |
| Has not had asthma                 | 91.9 (47,686)        | 88.4 (309,729)                       | 88.5 (107,495)                                   |

Note: For participants with complete data in fully adjusted model for age, sex, BMI, income, education, smoking status, passive smoking exposure. —, UK Biobank wheeze prevalence shown in second data column and shortness of breath prevalence shown in third data column; BMI, body mass index; SD, standard deviation.

Only 35% of UK Biobank subjects had data for shortness of breath because it was added to the baseline survey late in the recruitment phase (i.e., as of 2009).

1.05 per 5 μg/m³ (95% CI: 1.00, 1.10) and a statistically significant association with shortness-of-breath symptoms [OR = 1.28 per 5 μg/m³ (95% CI: 1.15, 1.42)]. Residential exposure to NO₂ also showed significant associations with both wheeze and shortness of breath.

Table 2. European Study of Cohorts for Air Pollution Effects (ESCAPE)-based annual average air pollution concentrations at place of residence for the periods 2009–2010 (Lifelines) and 2010–2011 (UK Biobank) in micrograms per cubic meter and correlation matrix.

| Pollutant  | Percentiles | Correlation coefficients (r) | NO2 |
|------------|-------------|------------------------------|-----|
|            | 5th         | 50th                         | 95th | PM2.5 | PM10 | PMcoarse |
| PM2.5      | Pooled      | 10.72 ± 0.98                | 9.22 | 10.66 | 12.39 | —       | —       | —       |
| Lifelines  | 15.55 ± 0.36| 15.27                        | 15.45 | 16.25 | —     | 1       | 0.78    | 0.37    | 0.42    |
| UK Biobank | 9.95 ± 1.04 | 8.26                        | 9.90 | 11.78 | 1     | 0.53    | 0.21    | 0.87    |
| PM10       | Pooled      | 17.29 ± 1.78                | 14.55 | 17.06 | 20.92 | —       | —       | —       | —       |
| Lifelines  | 24.23 ± 0.68| 23.73                        | 23.95 | 25.49 | 1     | 0.87    | 0.74    | 0.37    | 0.42    |
| UK Biobank | 16.18 ± 1.90| 13.08                        | 16.00 | 20.19 | 1     | 0.81    | 0.50    | —       | —       |
| PMcoarse   | Pooled      | 6.73 ± 0.85                 | 6.00 | 6.43  | 9.12  | —       | —       | —       |
| Lifelines  | 8.68 ± 0.46 | 8.29                        | 8.52 | 9.60  | 1     | 0.77    | 0.37    | 0.42    |
| UK Biobank | 6.41 ± 0.90 | 5.63                        | 6.10 | 9.04  | 1     | 0.19    | 0.37    | 0.42    |
| NO2        | Pooled      | 25.45 ± 7.13                | 14.40 | 24.35 | 36.93 | —       | —       | —       |
| Lifelines  | 16.48 ± 3.78| 11.93                        | 15.68 | 23.38 | 1     | 0.77    | 0.37    | 0.42    |
| UK Biobank | 26.28 ± 5.73| 14.79                        | 25.74 | 39.09 | 1     | 0.19    | 0.37    | 0.42    |

Note: For participants with complete data for wheeze, age, BMI, income, education, smoking status, passive smoking exposure (Lifelines n = 52,062; UK Biobank n = 325,892). —, pooled correlation coefficients not performed; BMI, body mass index; NO2, nitrogen dioxide; PM2.5, particulate matter with aerodynamic diameter <2.5 μm; PM10, particulate matter with aerodynamic diameter 2.5–10 μm; SD, standard deviation.

In two-pollutant models, the association between PM2.5 and wheeze was robust to adjustment for NO₂, but the association between PM2.5 and shortness of breath was not (Table 3). Associations of PM10 and PMcoarse were attenuated to the null for both outcomes when adjusted for NO₂. The association between
Table 3. Logistic regression model estimates of associations between annual average air pollution exposures at the baseline residence and respiratory symptoms for fully adjusted single-pollutant models (by cohort and pooled) and for pooled two-pollutant models.

| Exposure and outcome | UK Biobank | Lifelines | Adjusted for PM10 | Adjusted for NO2 |
|----------------------|------------|----------|------------------|------------------|
| **PM10** | OR (95% CI) | OR (95% CI) | OR (95% CI) | OR (95% CI) |
| **PM2.5** | OR (95% CI) | OR (95% CI) | OR (95% CI) | OR (95% CI) |
| **PMcoarse** | OR (95% CI) | OR (95% CI) | OR (95% CI) | OR (95% CI) |
| **NO2** | OR (95% CI) | OR (95% CI) | OR (95% CI) | OR (95% CI) |

Discussion

In this pooled study, we found statistically significant associations between mean annual PM and NO2 exposure and self-reported wheeze and shortness-of-breath symptoms. The strongest estimated associations were between fine particulate matter and self-reported shortness-of-breath symptoms, although there was little evidence of an association after adjustment for NO2. The use of large biobanks resulted in a large sample size that provided good statistical power to explore potentially vulnerable population subgroups. Our estimates suggested stronger associations of PM2.5 with both wheeze and shortness-of-breath symptoms, and of PMcoarse with shortness of breath, among participants with lower incomes compared with those with higher incomes. Associations between PM2.5 and wheeze were significantly higher among older versus younger participants, past or current smokers versus never smokers, and people without versus with asthma. In addition, PMcoarse was associated with shortness of breath among nonobese participants, but not among obese participants.

A previous longitudinal analysis of data from six European cohorts that also used ESCAPE exposure estimates reported positive but nonsignificant associations of NO2, nitrogen oxides (NOx), and PM exposure metrics with asthma incidence (Jacquemin et al. 2015), and another analysis based on data from four of these six cohorts reported nonsignificant positive NO2 and shortness of breath was not affected by adjustment for PM2.5, but NO2 was not associated with wheeze after adjustment (Table 3). Associations between pollutant exposure at residence and wheeze and shortness of breath were generally unchanged when restricting analyses to participants who lived at the same address for ≥5 y and ≥10 y (see Table S7).

Figures 2 and 3 outline pooled subgroup analyses for exposure to PM2.5 and PMcoarse, respectively. PM2.5 exposure was not associated with wheeze in higher-income individuals [OR = 1.02 (95% CI: 0.96, 1.08) per 5 μg/m3] but was associated with wheeze in lower-income participants [OR = 1.30 (95% CI: 1.22, 1.38); p-interaction <0.001]. Moreover, the association between shortness of breath and a 5-μg/m3 increment in PM2.5 was significantly higher among individuals from lower-income households [adjusted OR = 1.73 (95% CI: 1.52, 1.97)] than among higher-income individuals [adjusted OR = 1.31 (95% CI: 1.11, 1.55); p-interaction = 0.005]. Age, smoking status, and asthma status also modified associations between PM2.5 and wheeze symptoms, with slightly stronger associations for those ≥65 years of age (vs. <65; p-interaction = 0.004), past or current smokers (vs. never smokers; p-interaction <0.001) and people without asthma (vs. those with asthma; p-interaction = 0.005). With regard to PM2.5, the association between PMcoarse and shortness of breath was stronger for lower-income participants than for higher-income participants [OR = 1.41 (95% CI: 1.24, 1.61)] vs. 1.09 (95% CI: 0.91, 1.30); p-interaction = 0.009] although there was no clear difference in the association between PMcoarse and wheeze according to income. Lastly, PMcoarse was associated with shortness of breath among nonobese participants [OR = 1.41 (95% CI: 1.23, 1.63)] but not among obese participants [OR = 1.02 (95% CI: 0.87, 1.20); p-interaction = 0.002]. Both PM10 and NO2 were associated with wheeze in lower-income individuals but not in higher-income participants (see Figures S1 and S2). Associations between NO2 (but not PM10) and shortness of breath were also significantly stronger for lower-income individuals than for higher-income participants (see Figures S1 and S2). Lastly, interaction for obesity status was observed for PM10 and NO2 associations with shortness of breath, with nonobese participants showing significantly stronger associations.
Associations of NO₂, NOₓ, and PM₁₀ with COPD prevalence and incidence (Schikowski et al. 2014). More recently, significant positive associations were reported for PM₁₀ and NO₂ exposures (estimated using ESCAPE models) with lifetime asthma prevalence (a self-reported history of ever having had asthma) and current asthma (a self-reported history of asthma plus current or recent use of asthma medications) in a combined analysis of data from UK Biobank, Lifelines, and the Norwegian Nord-Trøndelag Health Study (HUNT3) cohorts (Cai et al. 2017). Associations between air pollutants and shortness of breath were stronger than corresponding associations with wheeze in our study population, which might be consistent with cardiovascular effects (Ebi-Kryston 1988) rather than (or in addition to) respiratory effects. A previous study of residential ambient air pollutant concentrations in Swiss adults reported associations with the prevalence of breathlessness, but not wheeze (Zemp et al. 1999), and adults living along busy streets in Netherlands were more likely to report shortness of breath during walking, but not wheeze, than adults living along low-traffic streets in the same neighborhoods (Oosterlee et al. 1996). However, these findings may also represent reporting misclassification: although individuals diagnosed with asthma are likely to recognize the symptom of wheeze, the rest of the population might not and

| Sex       | Number of participants | Odds ratio (95% CI) | Interaction P-value |
|-----------|------------------------|---------------------|---------------------|
| Males     | 176854                 | 1.16 [1.09 - 1.23]  | .451                |
|           | 79947                  | 1.57 [1.34 - 1.85]  | .867                |
|           | 201100                 | 1.15 [1.08 - 1.22]  | .751                |
|           | 93613                  | 1.62 [1.42 - 1.85]  |                     |
| Females   | 318950                 | 1.14 [1.09 - 1.19]  | .004                |
|           | 148873                 | 1.68 [1.50 - 1.89]  | .802                |
|           | 59004                  | 1.29 [1.17 - 1.43]  |                     |
|           | 24687                  | 1.61 [1.28 - 2.01]  |                     |
| Age       |                        |                     |                     |
| Under 65  | 173657                 | 1.30 [1.22 - 1.38]  | <.001               |
|           | 80101                  | 1.73 [1.52 - 1.97]  | .005                |
|           | 204297                 | 1.02 [0.96 - 1.08]  |                     |
|           | 93459                  | 1.31 [1.11 - 1.55]  |                     |
| 65 and older|                       |                     |                     |
| Income    |                        |                     |                     |
| Low       | 318950                 | 1.43 [1.37 - 1.50]  | .751                |
|           | 148873                 | 2.25 [2.15 - 2.35]  | .802                |
|           | 59004                  | 2.95 [2.77 - 3.13]  |                     |
|           | 24687                  | 3.19 [2.99 - 3.40]  |                     |
| High      |                        |                     |                     |
| Obesity status |                  |                     |                     |
| Non-obese | 318950                 | 1.13 [1.07 - 1.19]  | .068                |
|           | 148873                 | 1.68 [1.56 - 1.80]  | .100                |
|           | 59004                  | 2.02 [1.86 - 2.19]  |                     |
|           | 24687                  | 2.35 [2.17 - 2.54]  |                     |
| Obese     |                        |                     |                     |
| Smoking Status |                |                     |                     |
| Never smoked | 215185               | 1.08 [1.02 - 1.15]  | <.001               |
|            | 95568                  | 1.71 [1.48 - 1.97]  | .233                |
|            | 162769                 | 1.25 [1.18 - 1.33]  |                     |
|            | 77992                  | 1.49 [1.29 - 1.73]  |                     |
| Past or current smoker |            |                     |                     |
| Asthma status |                  |                     |                     |
| Non-asthmatic | 335245               | 1.19 [1.13 - 1.25]  | .005                |
|            | 155134                 | 1.58 [1.41 - 1.78]  | .857                |
|            | 42316                  | 1.11 [1.00 - 1.23]  |                     |
| Asthmatic  | 18163                  | 1.67 [1.34 - 2.08]  |                     |

Figure 2. Adjusted odds ratios [and 95% confidence intervals (CIs)] for respiratory symptoms in association with a 5-µg/m³ increase in ambient particulate matter with aerodynamic diameter <2.5 µm (PM₂.₅) at participant residences among population subgroups based on pooled data from the Lifelines and UK Biobank cohorts. Logistic regression model adjusted for age (continuous), sex, body mass index (BMI) (normal, overweight, or obese), household income (annual net income ≤ or > the country-specific mean for 2010), education level (≤ secondary or postsecondary), smoking status (never, former, or current), passive smoking exposure (none/any), and cohort (Lifelines/UK Biobank). Interaction p-values are Wald p-values for product interaction terms between air pollutants and stratification variables. Non-italic p-values (top), wheeze symptoms; Italic p-values (bottom), shortness-of-breath symptoms.
therefore might label the symptom as shortness of breath. Wheeze is not always well recognized in the general population. In the seminal International Study of Asthma and Allergies in Childhood (ISAAC), many countries used a video to demonstrate wheezing symptoms without using the word wheezing and asked if the child had these symptoms, as well as using a written questionnaire using the word "wheezing." Results comparing prevalence tertiles determined using the video versus those determined using the written questionnaire were discordant in a third of centers (Beasley 1998).

Our study provides evidence in line with previous reports that lower household income may be associated with greater impact of particulate matter pollution on respiratory health (Wheeler and Ben-Shlomo 2005), which is of concern because lower-SES individuals might be at increased risk for certain health effects of air pollution. As shown in a recent neighborhood-level study conducted in the same countries as these cohorts (i.e., Netherlands and the United Kingdom), the highest particulate matter and NO2 concentrations were consistently found in the most deprived areas in both countries (Fecht et al. 2015). Lower-SES individuals might be more susceptible to health effects of air pollution exposures because of coexposures to other environmental stressors (e.g., housing conditions, occupational exposures) or because of comorbid conditions related to reduced access to health care or to lifestyle, diet, and other factors (Lipfert 2004).

A review of factors associated with increased susceptibility to the health effects of air pollution suggested that older individuals and people with asthma are at increased risk (Sacks et al. 2011), consistent with our findings. Although some previous studies have reported evidence suggesting that obese people are more susceptible to health effects of air pollution than nonobese people (Sacks et al. 2011), we estimated stronger associations between PMcoarse, PM10, and NO2 and shortness of breath among nonobese individuals than among obese individuals in our study population. A subjective increase in shortness of breath among obese individuals might largely be explained by decreased respiratory

---

### Table: Adjusted odds ratios and 95% confidence intervals for respiratory symptoms in association with exposure to particulate matter

|                        | Number of participants | Odds ratio (95% CI) | Interaction P-value |
|------------------------|------------------------|---------------------|---------------------|
| **Sex**                |                        |                     |                     |
| Males                  | 17854                  | 1.01 [0.94 - 1.08]  | .220                |
| Females                | 79947                  | 1.31 [1.11 - 1.54]  | .803                |
|                        | 20110                  | 1.08 [1.01 - 1.16]  |                     |
|                        | 93613                  | 1.27 [1.10 - 1.45]  |                     |
| **Age**                |                        |                     |                     |
| Under 65               | 31865                  | 1.04 [0.99 - 1.10]  | .637                |
| 65 and older           | 14887                  | 1.35 [1.20 - 1.51]  | .258                |
|                        | 59004                  | 1.07 [0.95 - 1.20]  |                     |
|                        | 24687                  | 1.15 [0.91 - 1.46]  |                     |
| **Income**             |                        |                     |                     |
| Low                    | 173657                 | 1.07 [1.00 - 1.15]  | .443                |
|                        | 80101                  | 1.41 [1.24 - 1.61]  | .009                |
|                        | 204297                 | 1.03 [0.96 - 1.10]  |                     |
|                        | 93459                  | 1.09 [0.91 - 1.30]  |                     |
| High                   |                        |                     |                     |
| Non-obese              | 293097                 | 1.02 [0.96 - 1.08]  | .184                |
|                        | 137253                 | 1.41 [1.23 - 1.63]  | .002                |
|                        | 84857                  | 1.09 [1.00 - 1.19]  |                     |
|                        | 36307                  | 1.02 [0.87 - 1.20]  |                     |
| Obese                  |                        |                     |                     |
| Never smoked           | 215185                 | 1.04 [0.97 - 1.11]  | .526                |
|                        | 95568                  | 1.28 [1.11 - 1.48]  | .820                |
| Past or current smoker | 162789                 | 1.06 [0.99 - 1.14]  |                     |
|                        | 77992                  | 1.28 [1.10 - 1.49]  |                     |
| **Asthma status**      |                        |                     |                     |
| Non-asthmatic          | 335245                 | 1.05 [0.98 - 1.11]  | .940                |
|                        | 155134                 | 1.25 [1.11 - 1.41]  | .240                |
| Asthmatic              | 42316                  | 1.07 [0.95 - 1.20]  |                     |
|                        | 18163                  | 1.48 [1.18 - 1.85]  |                     |

Figure 3. Adjusted odds ratios (and 95% confidence intervals [CI]) for respiratory symptoms in association with a 5-µg/m3 increase in ambient particulate matter with aerodynamic diameter 2.5–10 µm (PMcoarse at participant residences among population subgroups based on pooled data from the Lifelines and UK Biobank cohorts. Logistic regression model adjusted for age (continuous), sex, body mass index (BMI) (normal, overweight, or obese), household income (annual net income ≤ or > the country-specific mean for 2010), education level (≤ secondary or postsecondary), smoking status (never, former, or current), passive smoking exposure (none/any), and cohort (Lifelines/UK Biobank). Interaction p-values are Wald p-values for product interaction terms between air pollutants and stratification variables. Non-italic p-values (top), wheeze symptoms; Italic p-values (bottom), shortness-of-breath symptoms.
ments used by each cohort, the way in which participants addressed. A study that compared exposures estimated with and without accounting for daily mobility in the metropolitan area and in southern California found that estimates that did not account for mobility were biased toward the null, with greater bias when the spatial variability of NO2, and the association between NO2 and shortness of breath was robust to adjustment for PM2.5, whereas other associations were attenuated to the null after adjustment for a second pollutant. However, the two-pollutant model results should be interpreted with caution because of the different sample sizes, the high correlation between pollutants, and the different pollutant correlation structures across the cohorts included in this study.

Fine particulates (i.e., PM2.5) have been reported to have stronger associations with respiratory or cardiovascular mortality relative to coarse (i.e., PM2.5–10) particulates (Brunekreef and Forsberg 2005; Faustini et al. 2014). This difference might be explained by a wider diffusion of fine particulates in the respiratory tract and by a stronger systemic inflammatory response. However, PMcoarse has been shown to have as strong or stronger short-term effects on COPD, asthma, and respiratory hospital admissions (Brunekreef and Forsberg 2005). For wheeze symptoms, we found consistently stronger associations with PM2.5 than with PMcoarse exposure in pooled and study-specific analyses. The association between PMcoarse and shortness of breath was stronger than the association between PM2.5 and shortness of breath in the Lifelines cohort, where participants were asked if they had ever had an attack of shortness of breath during the day when they were at rest. In contrast, the association between PM2.5 and shortness of breath was stronger than the association between PMcoarse and shortness of breath in the UK Biobank cohort, where participants were asked if they had ever had shortness of breath while walking. These divergent results might reflect differences between the study populations in the composition and sources of coarse particles, the data collection instruments used by each cohort, the way in which participants interpreted the questions about shortness of breath, or other factors (Giersing et al. 2010; Guillemin et al. 1993). Further, whereas the UK Biobank questionnaire collected period-specific wheeze prevalence (i.e., in the last year), the lack of a specified period in the Lifelines assessment item might have led to recall bias. To help overcome such challenges in the future, new studies should make use of internationally validated questionnaires whenever possible. Implementing standardized questionnaires on asthma and COPD respiratory symptoms and risk factors in epidemiological studies would greatly facilitate comparing and pooling data (Pistelli and Maio 2014).

A few limitations with regard to exposure assessment should be noted. First, because no information on study participants’ work addresses or time-activity patterns was available in either cohort, air pollution exposure was only estimated at the home address. A study that compared exposures estimated with and without accounting for daily mobility in the metropolitan Vancouver, British Columbia area and in southern California found that estimates that did not account for mobility were biased toward the null, with greater bias when the spatial variability of pollution was higher (Setton et al. 2011). Second, applying the ESCAPE southeast England LUR models to the larger UK Biobank study area could have led to exposure measurement error. However, LUR models provided good predictions of historical NO2 concentrations across the United Kingdom ($R^2 = 0.67$) and good predictions of PM10 concentrations within a 400-km radius of the initial ESCAPE study area ($R^2 = 0.53$) (Gulliver and de Hoogh 2015).

Pooled analyses would have been more strongly influenced by data from the UK Biobank cohort than from the Lifelines cohort because of the larger number of UK Biobank observations included in the analyses. Further, although only 45% of total Lifelines participants had been attributed air pollution exposure estimates, we do not expect selection bias to have been a major problem because participant residences were geocoded in chronological order (i.e., in the order of recruitment) rather than according to any given participant characteristic. Residual confounding, exposure and outcome misclassification, and other potential sources of bias are also possible. Finally, given the cross-sectional nature of our analyses, we cannot confirm the temporal relationship between air pollution exposures and the respiratory symptoms in our study population.

**Conclusion**

In conclusion, the findings from our cross-sectional analysis of data from adult cohorts in the United Kingdom and Netherlands add to existing evidence of the adverse effects of ambient air pollution on respiratory health in adults. In addition, differences in associations among population subgroups may reflect differences in susceptibility according to age, SES, and other factors.

**Acknowledgments**

The authors wish to acknowledge the services of the staff at UK Biobank and the Lifelines Cohort Study, the contributing research centers delivering data on these two cohorts, and all the study participants. We also acknowledge the ESCAPE project for their work in developing the air pollution exposure assessments used in our study. Lastly, we would like to thank Y. Marcon (Research Institute of the McGill University Health Centre, Montreal, Canada), P. Burton and D. Avraam (School of Social and Community Medicine, Bristol, UK), and C. Pang (University Medical Center Groningen, Groningen, Netherlands) for their support in implementing DataSHIELD federated analyses. Open-source software and the methodology used to harmonize and analyze data across cohorts were developed by Maelstrom Research (www.maelstrom-research.org).

The research leading to these results received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 261433 [Biobank Standardisation and Harmonisation for Research Excellence in the European Union (BioSHaRE-EU)]. Y.C. acknowledges support from the Early-Career Research Fellowship awarded by the UK Medical Research Council – Public Health England Centre for Environment and Health (Grant number MR/M501669/1).

**References**

Abramson M, Matheson M, Wharton C, Sim M. 2002. Prevalence of respiratory symptoms related to chronic obstructive pulmonary disease and asthma among middle aged and older adults. Respiratory 7(4):325–331, PMID: 12421240, https://doi.org/10.1093/1448-1943.2002.00408.x.

Allen N, Sudlow C, Downey P, Peasman T, Danesh J, Elliott P, et al. 2012. UK Biobank: current status and what it means for epidemiology. Health Policy Technol 1(3):123–126, https://doi.org/10.1016/j.hjht.2012.07.003.

Bayer-Oglesby L, Schindler C, Hazenkamp-von Arx ME, Braun-Fahrländer C, Keidel D, Rapp R, et al. 2006. Living near main streets and respiratory symptoms in adults: The Swiss Cohort Study on Air Pollution and Lung Diseases in Adults. Am J Epidemiol 164(12):1198–1198, PMID: 17026894, https://doi.org/10.1093/aje/kw308.

Beasley R. 1998. Worldwide variation in prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and atopic eczema: ISAAC. Lancet 351(9111):1225–1232, PMID: 9643741, https://doi.org/10.1016/S0140-6736(97)07302-9.

Beelen R, Hoek G, Vienneau D, Eeftens M, Dimakopoulou K, Pedelis K, et al. 2013. Development of NO2 and NOX land use regression models for estimating air pollution exposure in 36 study areas in Europe – the ESCAPE project. Atmos Environ 72:10–22, https://doi.org/10.1016/j.atmosenv.2012.03.037.

Bozkurt B, Mann DL. 2003. Shortness of breath. Circulation 108(2):e11–e13, PMID: 12680984, https://doi.org/10.1161/01.CIR.0000075956.36340.78.
