Long-term Associations Between Time-varying Exposure to Ambient $PM_{2.5}$ and Mortality

An Analysis of the UK Biobank

Jacopo Vanoli,[a] Arturo de la Cruz Libardi,[a] Francesco Sera,[a,b] Massimo Stafoggia,[c] Pierre Masselot,[a] Malcolm N. Mistry,[a,d] Sanjay Rajagopalan,[e,f] Jennifer K. Quint,[g] Chris Fook Sheng Ng,[h,i] Lina Madaniyazi,[i] and Antonio Gasparrini[a]

Downloaded from http://journals.lww.com/epidem by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8K2+Ya6H515kE= on 11/26/2024

Background: Evidence for long-term mortality risks of $PM_{2.5}$ comes mostly from large administrative studies with incomplete individual information and limited exposure definitions. Here we assess $PM_{2.5}$–mortality associations in the UK Biobank cohort using detailed information on confounders and exposure.

Methods: We reconstructed detailed exposure histories for 498,090 subjects by linking residential data with high-resolution $PM_{2.5}$ concentrations from spatiotemporal machine-learning models. We split the time-to-event data and assigned yearly exposures over a lag window of 8 years. We fitted Cox proportional hazard models with time-varying exposure controlling for contextual- and individual-level factors, as well as trends. In secondary analyses, we inspected the lag structure using distributed lag models and compared results with alternative exposure sources and definitions.

Results: In fully adjusted models, an increase of 10 μg/m³ in $PM_{2.5}$ was associated with hazard ratios of 1.27 (95% confidence interval: 1.06, 1.53) for all-cause, 1.24 (1.03, 1.50) for nonaccidental, 2.07 (1.04, 4.10) for respiratory, and 1.66 (0.86, 3.19) for lung cancer mortality. We found no evidence of association with cardiovascular deaths (hazard ratio = 0.88, 95% confidence interval: 0.59, 1.31). We identified strong confounding by both contextual- and individual-level lifestyle factors. The distributed lag analysis suggested differences in relevant exposure windows across mortality causes. Using more informative exposure summaries and sources resulted in higher risk estimates.

Conclusions: We found associations of long-term $PM_{2.5}$ exposure with all-cause, nonaccidental, respiratory, and lung cancer mortality, but not with cardiovascular mortality. This study benefits from finely reconstructed time-varying exposures and extensive control for confounding, further supporting a plausible causal link between long-term $PM_{2.5}$ and mortality.

Keywords: Cohort study; Distributed lag models; Machine learning; Mortality; Particulate matter; UK Biobank

(Epidemiology 2025;36: 1–10)

Chronic exposure to fine particulate matter ($PM_{2.5}$) is recognized as a major risk factor for human health. The epidemiologic literature has focused particularly on nonaccidental mortality, for which now extensive evidence supports the existence of a causal link.[1–3] However, strong associations have also been found for cause-specific outcomes, including cardiovascular, respiratory, and lung cancer mortality,[4–6] with recent evidence extending the link to other outcomes such as metabolic and kidney diseases as well as neurological disorders.[7–9] These associations were shown to persist at low $PM_{2.5}$ concentrations (below 10–12 μg/m³), prompting the World Health Organization to revise the recommended annual limit downward to 5 μg/m³.

Submitted December 16, 2023; accepted September 23, 2024

From the [a]Environment & Health Modelling (EHM) Lab, Department of Public Health Environments and Society, London School of Hygiene & Tropical Medicine, London, United Kingdom; [b]Department of Statistics, Computer Science and Applications "G. Parenti," University of Florence, Florence, Italy; [c]Department of Epidemiology, Lazio Region Health Service, ASL Roma 1, Rome, Italy; [d]Department of Economics, Ca' Foscari University of Venice, Venice, Italy; [e]Harrington Heart and Vascular Institute, University Hospitals Cleveland Medical Center, Cleveland, OH; [f]School of Medicine, Case Western Reserve University, Cleveland, OH; [g]School of Public Health, Imperial College London, London, United Kingdom; [h]Department of Global Health Policy, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; and [i]School of Tropical Medicine and Global Health, Nagasaki University, Nagasaki, Japan.

Supported by the Nagasaki University "Doctoral Program for World-leading Innovative and Smart Education" for Global Health (WISE), KENKYU SHIDO KEIHI ("the Research Grant"), and KYOIKU KENKYU SHIEN KEIHI ("the Stipend"), and by the Medical Research Council UK (grant ID: MR/Y003330/1).

Disclosure: The authors report no conflicts of interest.

The original full R code used to perform the analysis is available in a GitHub repo (https://github.com/gasparrini/UKB-pm25deathlong), together with a simpler reproducible code illustrating a tutorial example using synthetic datasets that resemble the original UK Biobank data. The actual database of the UK Biobank can be accessed via an application through the UK Biobank online Access Management System (AMS) (https://www.ukbiobank.ac.uk/). The $PM_{2.5}$ exposure dataset was expressly linked in this analysis and will be fully attached to the main database and released in the UK Biobank repository.

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

Correspondence: Antonio Gasparrini, London School of Hygiene & Tropical Medicine, 15-17 Tavistock Place WC1H 9SH, London, United Kingdom. E-mail: antonio.gasparrini@lshtm.ac.uk.

Copyright © 2024 Wolters Kluwer Health, Inc. All rights reserved.
ISSN: 1044-3983/25/361-110
DOI: 10.1097/EDE.0000000000001796

Copyright © 2024 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

Downloaded from http://journals.lww.com/epidem by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8K2+Ya6H515kE= on 11/26/2024

The large number of studies on the long-term impact of $PM_{2.5}$ on mortality have been summarized by various meta-analyses.[10,11] The main evidence, in terms of sample size and population representativeness, comes from large administrative cohorts in North America, including the historical Harvard Six Cities and the American Cancer Society cohorts,[12] and more recently the US Medicare[13] and the Canadian Census Health and Environment Cohorts (CanCHEC),[14] as well as multicenter studies conducted in Europe, for instance the European Study of Cohorts for Air Pollution Effects (ESCAPE) and Effects of Low-Level Air Pollution: A Study in Europe (ELAPSE).[15–17]

Despite this evidence base, there are limitations. First, most of the large studies are based on administrative cohorts with partial information on individual-level characteristics.[13,14,17] Specifically, such analyses ensured control of socioeconomic and demographic attributes such as age, sex, income, and education but lacked information on lifestyle and other health-related factors. These characteristics, including smoking, alcohol consumption, and physical exercise, as well as body mass index, are among the strongest risk factors for mortality and have a strong potential for confounding.

Second, a critical feature of these studies of long-term $PM_{2.5}$ and mortality is the exposure assignment process. Recent cohort studies took advantage of state-of-the-art exposure models that provide fine spatiotemporal resolution and accuracy for retrospective assignment of time-varying exposures.[18,19] However, most analyses relied on simple exposure definitions, for instance using same-year $PM_{2.5}$ levels,[13] assigning time-constant values,[15–17] or using single predetermined exposure windows.[14] Previous studies have stressed the sensitivity of the estimates to the spatial and temporal scale of the exposure assignment,[20] and a more informative definition of the exposure–risk summaries, accounting for time-varying levels and individual residential histories, can lead to improved health risk estimates and knowledge of etiologic mechanisms.

Third, both meta-analyses and pooled studies indicated a degree of heterogeneity in the estimates of health risks associated with long-term exposure to $PM_{2.5}$.[10,11,17] While part of it can be explained by differences in exposure assessment as well as residual confounding, real variations in risks can also be related to particulate composition or underlying vulnerability of populations. Country and region-specific assessments are therefore needed for reliable health risk estimation at the local level. However, evidence from the United Kingdom is scarce, with only a few individual-level studies conducted using administrative databases and population-based cohorts.[21–23]

In this contribution, we assessed long-term risks for all-cause and cause-specific mortality associated with exposure to low levels of $PM_{2.5}$ in the UK Biobank (UKB) cohort. The study benefits from the analysis of a large cohort database with detailed individual characteristics, the use of detailed exposure histories assigned from high-resolution machine-learning (ML) models and accounting for residential mobility, and the application of state-of-the-art analytic methods to investigate temporal relationships.

METHODS

The UK Biobank

The UKB is a British prospective cohort study that has enrolled more than half a million participants aged 40–69 years between 2006 and 2010.[24] At enrollment, the participants underwent a first in-person assessment in one of the 22 assessment centers located across Great Britain (England, Scotland, and Wales). The assessment consisted of multiple questionnaires collecting information on lifestyles and personal characteristics, as well as the collection of anthropometric measures and biologic samples. Participants are tracked directly through follow-up assessments involving subsets of subjects and indirectly through routine linkage with various administrative health databases including mortality, hospital episode statistics, cancer screening, and primary care. The residential data are available in the UKB database, including the exact dates participants moved to each residence and the related geocoded locations with 100 m rounding. The data were validated internally, and the mobility history was continuously updated through general practitioner registration or direct reporting by the participants. The cohort profile was described in previous publications.[24,25] Details regarding the UKB database can be found on the showcase website (https://biobank.ndph.ox.ac.uk/showcase/). Access to the UKB database was permitted through the UK Biobank online Access Management System (approved research number: 56431). Ethical approval was granted by the LSHTM Ethics Committee (ref: 26679).

Mortality Outcomes

At the time of enrollment, the subjects consented to access to a variety of personal information, including linked electronic health records and residential address location. The health data included, among other outcomes, mortality records routinely extracted and updated from the UK national registers. We defined mortality outcomes based on the date and primary cause of death included in the main dataset following the International Classification of Disease 10th (ICD-10) revision: specifically, we defined all-cause (ICD-10 codes: A00–U99), nonaccidental (A00–R99), cardiovascular (I00–I99), respiratory (J00–J99), and lung cancer (C34) mortality outcomes.

Exposure Assessment and Linkage

Existing $PM_{2.5}$ exposure data available in the UKB database consist of annual average levels for the year 2010 assigned at the residential address established at recruitment using a Europe-wide land-use regression model with a resolution of 100×100 m.[26] This exposure linkage covered the subset of participants living within 400 km of London (see the first section and eTable 1; http://links.lww.com/EDE/C185 in the eAppendix). To enable the assignment of time-varying exposures across

© 2024 Wolters Kluwer Health, Inc. All rights reserved.

Copyright © 2024 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

Downloaded from http://journals.lww.com/epidem by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8K2+Ya6H515kE= on 11/26/2024

the whole cohort, we used an improved exposure data source and the full residential history of each participant. The new exposure data source consisted of daily $PM_{2.5}$ levels predicted on a 1-km grid across the United Kingdom in the period 2003–2021 using a hybrid spatiotemporal ML model. The model used an ensemble of ML algorithms trained using ground monitor series and a series of spatial and spatiotemporal predictors, including outputs from emission-dispersion models, remote sensing satellite data, and land use and traffic variables, among others.[27] The model performance, assessed using monitor-based cross-validation, provided an overall coefficient of determination (R^2) of 0.81 on a daily scale.

The exposure assignment process is described in previous work[28] and briefly summarized here. The process was performed in two steps. First, we constructed daily exposure series for each home location by combining gridded exposure values using bilinear interpolation, a process that performs a weighted average of the four nearest 1×1 km cells using weights computed as the inverse distance of the location from the corresponding centroids. This approach allowed maximizing the exposure information while masking the original data, thus preventing backtracing of individual locations. Second, we composed the subject-specific exposure profiles by linking the daily series for corresponding periods determined by the residential history. We aggregated the daily subject-specific exposure data in annual series corresponding to calendar years to facilitate the model fitting, as described below.

Study Design and Statistical Analysis

We performed a time-to-event analysis based on an extended Cox proportional hazard model with time-varying predictors. Subjects were censored at the time of death, lost to follow-up, or the administrative end of follow-up (set here to 31 December 2021), whichever came first. We defined events as death for any cause or specific causes only. We specified the Cox model using calendar time as the temporal axis and stratified by assessment center, sex, and year of birth, thus ensuring a strong control for temporal trends and differential risks by age. We split the time-to-event data by calendar year (defining yearly subject/periods starting on 1 January) to assign exposure levels over a window of 8 years, consistently with previous studies.[14] Specifically, for each subperiod, we assigned the yearly averaged $PM_{2.5}$ from the previous calendar year (defined here as lag 0) to 8 years earlier (lag 7), so that exposure always preceded a mortality event. Given the need to define complete 8-year exposure histories and that the exposure data covered the period from 2003 on, the follow-up period effectively started on 1 January 2011 for all the subjects. We dropped from the analysis events that occurred before this date and subjects with incomplete exposure history. We assumed a linear relationship and, in the main analysis, we defined the exposure index as the average of $PM_{2.5}$ within four windows: 1 year (lag 0), 2 years (lag 0–1), 5 years (lag 0–4), and 8 years (lag 0–7). Each window was investigated in separate models.

We specified six confounder models with increasing control for individual and area-level covariates measured at baseline, all entered as linear or categorical terms. The selection of confounders was based on a directed acyclic graph that defined the assumed causal pathways (eFigure 1; http://links.lww.com/EDE/C185). Model 1 only included the matching variables used for stratification (assessment center, sex, and year of birth). Models 2 and 3 added control for socioeconomic status, using contextual- and individual-level socioeconomic variables, respectively. Specifically, Model 2 used the area-level Townsend deprivation index measured in 2010, while Model 3 included ethnic background, education level, household income, and employment status determined at recruitment. Model 4 featured all the variables in the previous models, while Model 5 also featured urban–rural classification and green space, two additional area-level variables rarely controlled for in published studies. Finally, Model 6 added a list of lifestyle-related factors rarely available in previous studies, including smoking status and intensity, alcohol intake, waist-to-hip ratio, physical activity (measured using the International Physical Activity Questionnaire scale), and living alone. Details on the variable definitions are provided in the supplementary appendix (first section and eTable 1; http://links.lww.com/EDE/C185 in the eAppendix).

In secondary analyses, we inspected the lag structure using distributed lag models (DLMs),[29] using the extended version for time-to-event data.[30] Specifically, we applied two different functions to specify the lag–response relationship in the fully adjusted model (Model 6): natural cubic splines with four degrees of freedom, and strata defining steps for lags 0, 1–2, 3–4, and 5–7.

We performed several sensitivity analyses. First, we replicated the main model using alternative exposure sources and definitions, specifically using time-constant indices using either the value previously assigned from the land-use regression model or the average for 2010 from our ML model.[26,27] In addition, we repeated the analysis truncating the follow-up to the pre-2020 or the post-2013 periods. The former was used to remove the years affected by the COVID-19 pandemic, and the latter to include a washout period in order to account for potential healthy volunteer and other selection biases, as recently recommended.[31]

Estimates of the associations were reported as hazard ratios (HRs) for all-cause and specific mortality outcomes per 10 $\mu g/m^3$ increments in $PM_{2.5}$, with 95% confidence intervals (CIs). Missing values in the baseline covariates were imputed using multiple imputation by chained equation, producing five imputed datasets, with estimates combined using Rubin's rule.[32]

The original full R code and a simpler reproducible version illustrating the analysis using synthetic datasets are available in a GitHub repo (https://github.com/gasparrini/UKB-pm25deathlong).

© 2024 Wolters Kluwer Health, Inc. All rights reserved.

Copyright © 2024 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

Downloaded from http://journals.lww.com/epidem by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8K2+Ya6H515kE= on 11/26/2024

RESULTS

Population and Exposure Characteristics

The original dataset included 502,381 individuals, from which 4291 (0.85%) participants were excluded due to (partially) missing exposure histories, providing a final cohort of 498,090 individuals. The participants were followed up for an average of 10.3 years, with a total of 5,117,660 person-years. During the follow-up, there were 33,817 deaths among the selected participants, with 31,791 for nonaccidental causes, including 6,904 for cardiovascular, 2,461 for respiratory outcomes, and 2,820 for lung cancer. Tables 1 and 2 report descriptive statistics for the selected cohort. The average age at recruitment was 56.5 years. The participants included 54.5% women, and it was overwhelmingly represented by people of White ethnicity (94.1%) and with high education (69.5% with a high school diploma or higher). Only 10.4% of the cohort were smokers at enrollment, most of them (65.0%) performed moderate or high physical activity, and only 8.0% consumed no alcohol. The majority (84.3%) lived in urban areas.

Figure 1 illustrates the box-and-whiskers plot with the distribution of the annual average exposure to $PM_{2.5}$ across the years. The plot shows that, across the study period, most of the participants were exposed to levels of air pollution that can be considered low (<15 μg/m³), and all lived in areas below the current UK air quality guideline level (25 μg/m³). Exposure levels show a noticeable temporal variation and a reduction from 2015 onwards, with most individuals living in areas with $PM_{2.5}$ levels below 10 μg/m³ in 2021. However, the majority of them were still exposed to concentrations above the current recommended limit by the World Health Organization (5 μg/m³).

TABLE 1. Descriptive Statistics (Mean with 5th–95th Percentile Range and Missing) for Continuous Baseline Characteristics[a] in the Selected UKB cohort (498,090 participants Recruited in the Period 2006–2010)

Age at baseline (years)	56.50 (42.00 to 68.00)
Missing (%)	0 (0.0%)
Waist-to-hip ratio	0.87 (0.73 to 1.02)
Missing (%)	2,165 (0.4%)
Smoking intensity (pack-years)	8.22 (0.00 to 40.62)
Missing (%)	77,321 (15.5%)
Townsend deprivation index	−1.30 (−5.05 to 4.92)
Missing (%)	616 (0.1%)
Green space (%)	45.01 (15.58 to 86.94)
Missing (%)	60,318 (12.1%)

[a]See the text and eTable 1; http://links.lww.com/EDE/C185 in the appendix for definitions and additional information.

Mortality Risks Associated with Long-term $PM_{2.5}$ Exposure

The estimates of the association between long-term exposure to $PM_{2.5}$ and mortality are reported in Table 3, with combinations of causes of death, exposure windows, and various levels of covariate adjustments. In the fully adjusted model (Model 6) and considering a cumulative exposure in the last 8 years (lag 0–7), an increase of 10 μg/m³ in $PM_{2.5}$

TABLE 2. Distributions of Categorical Baseline Characteristics[a] in the Selected UKB Cohort (498,090 Participants Recruited in the Period 2006–2010)

Sex, n (%)	Female	271,660 (54.5%)	Rural/urban classification, n (%)	Urban	420,043 (84.3%)
	Male	226,430 (45.5%)		Town/fringe	37,256 (7.5%)
	Missing (%)	0 (0.0%)		Village/rural	35,893 (7.2%)
Ethnic background, n (%)	White	468,508 (94.1%)		Missing (%)	4,898 (1.0%)
	Other	26,847 (5.4%)	Physical activity, n (%)	Low	75,255 (15.1%)
	Missing (%)	2,735 (0.5%)		Moderate	162,774 (32.7%)
Employment, n (%)	Employed	285,717 (57.4%)		High	161,041 (32.3%)
	Retired	164,864 (33.1%)		Missing (%)	99,020 (19.9%)
	Other	41,926 (8.4%)	Alcohol consumption, n (%)	Never	40,046 (8.0%)
	Missing (%)	5,583 (1.1%)		Special occasions only	57,388 (11.5%)
Education, n (%)	Low	83,964 (16.9%)		One to three times a month	55,424 (11.1%)
	Professional qualification	57,948 (11.6%)		Once or twice a week	128,297 (25.8%)
	High school diploma	186,136 (37.4%)		Three or four times a week	114,678 (23.0%)
	College/university degree	160,098 (32.1%)		Daily or almost daily	100,795 (20.2%)
	Missing (%)	9,944 (2.0%)		Missing (%)	1,462 (0.3%)
Income (£); n (%)	<18,000	95,735 (19.2%)	Smoking status, n (%)	Never	271,865 (54.6%)
	18,000–30,999	107,266 (21.5%)		Previous	171,301 (34.4%)
	31,000–51,999	110,145 (22.1%)		Current	52,037 (10.4%)
	52,000 to 100,000	85,892 (17.2%)		Missing (%)	2,887 (0.6%)
	>100,000	22,831 (4.6%)	Living alone, n (%)	No	402,043 (80.7%)
	Missing (%)	76,221 (15.3%)		Yes	92,937 (18.7%)
				Missing (%)	3,110 (0.6%)

[a]See the text and eTable 1; http://links.lww.com/EDE/C185 in the appendix for definitions and additional information.

© 2024 Wolters Kluwer Health, Inc. All rights reserved.

Copyright © 2024 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

Downloaded from http://journals.lww.com/epidem by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8K2+Ya6H515kE= on 11/26/2024

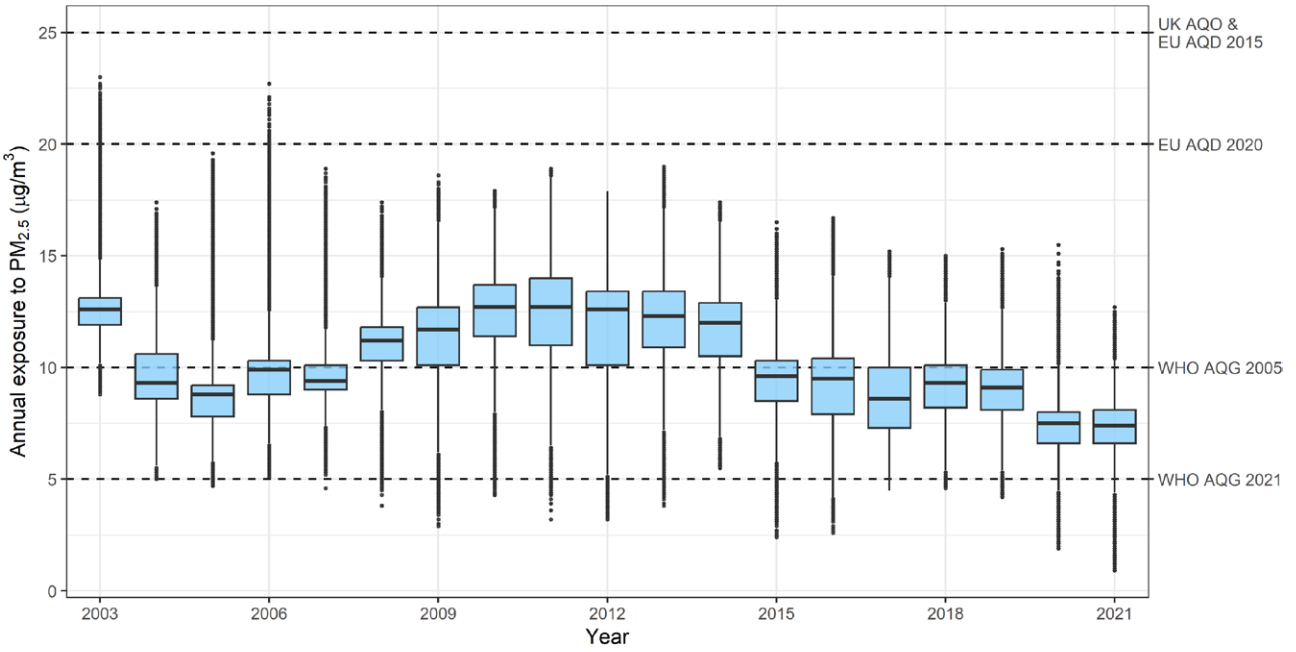

FIGURE 1. Distribution of annual average exposure to $PM_{2.5}$ across participants of the UKB cohort in the period 2003–2021, with limits corresponding to air quality guidelines (AQG) and air quality directives (AQD) from the World Health Organization (WHO), European Union (EU), and the United Kingdom (UK).

was associated with an HR of 1.27 (95% CI: 1.06, 1.53) and 1.24 (95% CI: 1.03, 1.50) for all-cause and nonaccidental mortality, respectively. There was evidence of positive associations also with respiratory (HR = 2.07, 95% CI: 1.04, 4.10) and lung cancer (HR = 1.66, 95% CI: 0.86, 3.19) mortality, although with wider confidence intervals, due to the smaller number of events. Surprisingly, cardiovascular mortality was not associated with long-term exposure to $PM_{2.5}$ (HR = 0.88, 95% CI: 0.59–1.31). In general, the associations were positive across lags and their magnitude increased with longer lag windows, with the exception of respiratory mortality.

The comparison of the estimates across models with different covariate adjustments in Table 3 indicates strong confounding effects, more pronounced for area-level factors compared with individual-level characteristics. The HR for nonaccidental mortality and lag 0–7 dropped from 4.11 (95% CI: 3.50, 4.82) in the basic model with only stratifying variables (Model 1) to 1.43 (95% CI: 1.21, 1.69) when area-level deprivation was accounted for (Model 2). Adding individual-level socioeconomic variables (Model 4) had little effect, while the inclusion of urban–rural residential settings and area-level green space (Model 5) further reduced the risk to an HR of 1.31 (95% CI: 1.09, 1.59). Interestingly, the results suggested noticeable confounding effects from other individual-level covariates represented by lifestyle variables (Model 6) for most of the mortality causes, with the estimated risk decreasing further for all-cause, nonaccidental, and lung cancer mortality while remaining stable for respiratory causes.

Analysis of the Lag Structure

In secondary analyses, we investigated potential differences within the relevant exposure window by extending the Cox model with the simpler cumulative exposure through DLMs, specifically using natural splines and step functions. The results are reported in Figure 2, with the estimated lag–response relationships reporting the HR along lags for an increase of 10 μg/m³ in $PM_{2.5}$. It should be acknowledged immediately that the relatively low number of events and associated statistical power resulted in low precision and wide confidence intervals, preventing firm conclusions. However, the graphs provide interesting suggestions. For instance, the analysis of the lag–response curves for all-cause and nonaccidental mortality indicates positive associations with exposures experienced in the last year (lag 0) and at longer timeframes (lag 5–7). In contrast, respiratory and lung cancer mortality showed associations for different exposure windows, specifically lag 0–2 for the former and lag 3–5 for the latter, suggesting differential temporal relationships related to different pathophysiological processes. The estimates of the overall cumulative associations are close to the main model with a simple exposure definition assuming a constant effect across lags (eTable 2; http://links.lww.com/EDE/C185).

Sensitivity Analyses Using Different Exposure Indices and Follow-up Periods

In the first sensitivity analysis, we replicated the main model using alternative exposure summaries and sources. The

© 2024 Wolters Kluwer Health, Inc. All rights reserved.

Copyright © 2024 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

TABLE 3. Hazard Ratios (HRs, with 95% Confidence Intervals) of Different Mortality Causes Associated with an Increase of 10 μg/m³ in $PM_{2.5}$ in the UKB Cohort, for Combinations of Length of Exposure Windows and Confounding Control

	Lag	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
All causes (33,817 events)	0	3.12 (2.73, 3.56)	1.37 (1.19, 1.56)	2.20 (1.92, 2.51)	1.33 (1.16, 1.53)	1.25 (1.07, 1.45)	1.20 (1.03, 1.39)
	0–1	3.49 (3.04, 4.01)	1.40 (1.22, 1.62)	2.38 (2.07, 2.73)	1.37 (1.19, 1.58)	1.27 (1.09, 1.50)	1.21 (1.03, 1.43)
	0–4	3.96 (3.42, 4.58)	1.44 (1.24, 1.67)	2.60 (2.24, 3.02)	1.40 (1.21, 1.63)	1.30 (1.10, 1.55)	1.24 (1.04, 1.47)
	0–7	4.44 (3.80, 5.18)	1.47 (1.26, 1.73)	2.84 (2.43, 3.33)	1.45 (1.23, 1.70)	1.34 (1.12, 1.61)	1.27 (1.06, 1.53)
Nonaccidental (31,791 events)	0	2.95 (2.58, 3.39)	1.34 (1.17, 1.54)	2.11 (1.84, 2.43)	1.32 (1.14, 1.51)	1.24 (1.06, 1.44)	1.18 (1.01, 1.38)
	0–1	3.26 (2.82, 3.75)	1.37 (1.18, 1.58)	2.26 (1.95, 2.61)	1.34 (1.16, 1.55)	1.25 (1.06, 1.48)	1.19 (1.01, 1.40)
	0–4	3.66 (3.15, 4.26)	1.39 (1.19, 1.62)	2.45 (2.10, 2.85)	1.36 (1.17, 1.59)	1.27 (1.06, 1.51)	1.20 (1.01, 1.43)
	0–7	4.11 (3.50, 4.82)	1.43 (1.21, 1.69)	2.68 (2.28, 3.15)	1.41 (1.20, 1.67)	1.31 (1.09, 1.59)	1.24 (1.03, 1.50)
Cardiovascular (6904 events)	0	3.59 (2.67, 4.83)	1.15 (0.86, 1.55)	2.08 (1.54, 2.80)	1.07 (0.79, 1.44)	1.05 (0.76, 1.47)	1.02 (0.73, 1.41)
	0–1	3.99 (2.93, 5.43)	1.13 (0.83, 1.54)	2.19 (1.60, 2.99)	1.04 (0.76, 1.42)	1.02 (0.72, 1.45)	0.97 (0.69, 1.38)
	0–4	4.31 (3.12, 5.96)	1.06 (0.77, 1.47)	2.25 (1.62, 3.12)	0.98 (0.70, 1.36)	0.94 (0.64, 1.36)	0.90 (0.62, 1.30)
	0–7	4.80 (3.40, 6.77)	1.04 (0.73, 1.48)	2.41 (1.70, 3.42)	0.97 (0.68, 1.38)	0.92 (0.62, 1.38)	0.88 (0.59, 1.31)
Respiratory (2461 events)	0	12.23 (7.33, 20.41)	2.01 (1.21, 3.37)	5.67 (3.36, 9.57)	1.93 (1.14, 3.26)	2.25 (1.24, 4.09)	2.18 (1.21, 3.94)
	0–1	14.34 (8.48, 24.23)	2.04 (1.20, 3.46)	6.28 (3.67, 10.76)	1.95 (1.14, 3.36)	2.34 (1.25, 4.37)	2.23 (1.20, 4.16)
	0–4	17.09 (9.87, 29.58)	2.05 (1.17, 3.57)	7.00 (3.99, 12.28)	1.96 (1.11, 3.46)	2.36 (1.22, 4.56)	2.23 (1.16, 4.28)
	0–7	18.63 (10.45, 33.21)	1.93 (1.08, 3.47)	7.28 (4.04, 13.11)	1.88 (1.03, 3.41)	2.21 (1.11, 4.40)	2.07 (1.04, 4.10)
Lung cancer (2820 events)	0	8.27 (5.20, 13.15)	1.72 (1.08, 2.74)	5.01 (3.11, 8.05)	1.80 (1.12, 2.89)	1.77 (1.05, 3.01)	1.52 (0.90, 2.58)
	0–1	9.39 (5.80, 15.20)	1.67 (1.03, 2.72)	5.40 (3.29, 8.84)	1.74 (1.07, 2.86)	1.71 (0.98, 2.99)	1.42 (0.81, 2.49)
	0–4	13.13 (7.86, 21.93)	1.90 (1.13, 3.20)	7.13 (4.21, 12.07)	2.01 (1.18, 3.41)	2.02 (1.10, 3.70)	1.62 (0.89, 2.98)
	0–7	15.96 (9.20, 27.69)	1.91 (1.09, 3.35)	8.32 (4.73, 14.63)	2.07 (1.17, 3.66)	2.06 (1.08, 3.95)	1.66 (0.86, 3.19)

Model 1: stratifying variables only (assessment center, sex, and month of age); Model 2: Model 1 + area-level deprivation (Townsend deprivation index); Model 3: Model 1 + individual-level socioeconomic variables (ethnic background, education level, household income, and employment status); Model 4: Model 2 + Model 3; Model 5: Model 4 + urban–rural classification and greenspace; and Model 6: Model 5 + individual-level lifestyle factors (smoking status and pack/years, alcohol intake, physical activity, waist-to-hip ratio, and living alone).

Downloaded from http://journals.lww.com/epidem by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8K2+Ya6H515kE= on 11/26/2024

assessment included a subset of 457,925 UKB participants (91.9% of the cohort used in the main analysis) with linked $PM_{2.5}$ levels assigned from the original land use regression model and the new ML model.[26,27] The correlation between the two sources for the year 2010 is low, with a Pearson *r* coefficient of 0.25. The analysis of the $PM_{2.5}$ distributions, shown in eFigure 2; http://links.lww.com/EDE/C185, reveals a peculiar pattern in the original data from the land use regression model, with the $PM_{2.5}$ left-truncated at about 8 μg/m³, while the new ML model provides a wider exposure range and bimodal shape probably due to urban/rural residential locations.

Figure 3 illustrates the comparison of the HRs estimated for each mortality cause using the original time-varying annual $PM_{2.5}$ from the ML model, the fixed annual average in 2010 from the same model, and the corresponding value for the same year from the existing UKB data derived from the land-use regression model (see eTable 3; http://links.lww.com/EDE/C185 for the actual HR figures). The graph shows that the use of time-varying exposure summary results in generally higher HRs, although with wider confidence intervals.

Further sensitivity analyses assessed the sensitivity to the right and left truncation of the follow-up period, with results reported in eTable 4; http://links.lww.com/EDE/C185. The exclusion of the period affected by the COVID-19 pandemic (pre-2020) generated very similar results. In contrast, the inclusion of a washout period with the restriction of the follow-up to the post-2013 years resulted in a general increase of the HR for all the mortality causes.

DISCUSSION

This epidemiologic study assessed long-term associations between time-varying $PM_{2.5}$ exposure and mortality in the UK Biobank cohort. Results show increased risks of death for all causes as well as nonaccidental, respiratory, and lung cancer, but not for cardiovascular outcomes. The comparison of models with different levels of adjustment for confounders indicates the important role of contextual variables such as area-level deprivation and green space, but also the need to control for individual-level factors related to lifestyle characteristics. The analysis features the innovative use of statistical techniques to analyze relevant lag structures through the implementation of distributed lag models within Cox proportional hazard regression, with suggestions of possible differences in the relevant exposure windows across mortality outcomes.

Our risk associations are generally higher than those reported in meta-analyses as well as large-scale or pooled studies, both for nonaccidental mortality and other causes, especially respiratory and lung cancer.[10,11,13–17] Prior results for the United Kingdom are sparser and less consistent: an analysis of a national English cohort of patients registered to clinical practices reported an HR of 1.13 (95% CI: 1.00, 1.29) for all-cause mortality,[21] while a recent analysis of the UKB

© 2024 Wolters Kluwer Health, Inc. All rights reserved.

Copyright © 2024 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

Downloaded from http://journals.lww.com/epidem by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8K2+Ya6H515kE= on 11/26/2024

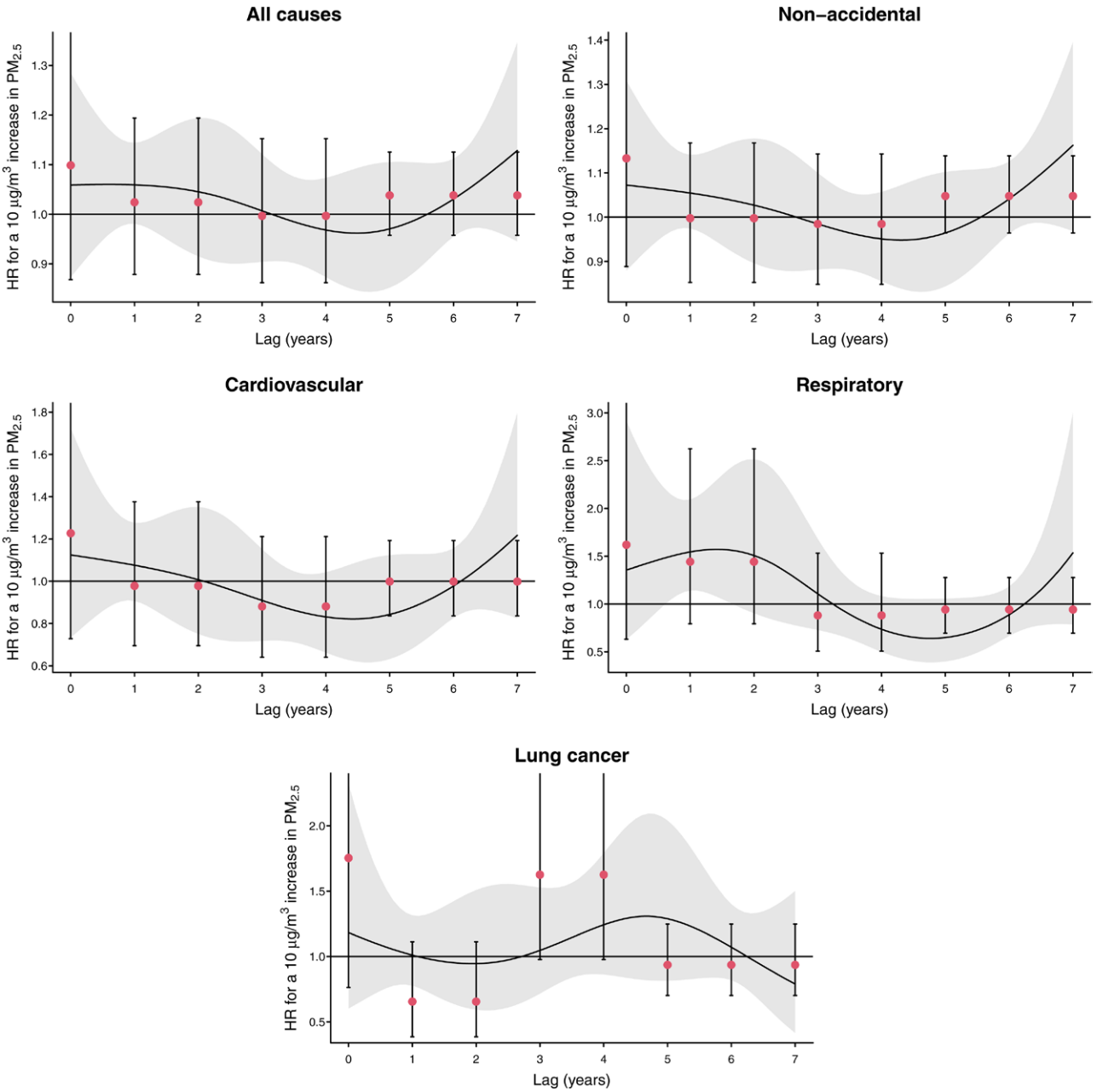

FIGURE 2. Lag–response relationships representing the hazard ratios (HRs, with 95% confidence intervals) of different mortality causes associated with an increase of 10 µg/m³ in $PM_{2.5}$ along an 8-year exposure window (lag 0–7), estimated by distributed lag models (DLM) with splines (line with 95% confidence intervals area) and step functions (points with 95% confidence intervals bars), respectively, using the fully adjusted confounder model (Model 6 in Table 3).

cohort found a higher risk of 1.27 (95% CI: 1.05, 1.55),[23] similar to that found by our study for all-cause mortality, although using time-constant exposure summaries and a less strict confounding control compared to our analysis. There can be multiple explanations for such differences. First, the use of more informative exposure data can have led to a reduction in exposure misclassification and a better ability to identify risk associations, as discussed later. Second, the risk could be steeper at the relatively low exposure ranges of the UKB cohort, as previously reported.[16,17,33] Interestingly, the higher size of the risk estimates using time-varying exposures and the findings about the confounding effects of lifestyle and other individual-level factors are consistent with other published analyses.[20,34]

In contrast, we found no evidence of an association with cardiovascular mortality. Such a link is in fact well established in the epidemiologic literature summarized earlier, although conflicting results were reported from the United

© 2024 Wolters Kluwer Health, Inc. All rights reserved.

Copyright © 2024 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

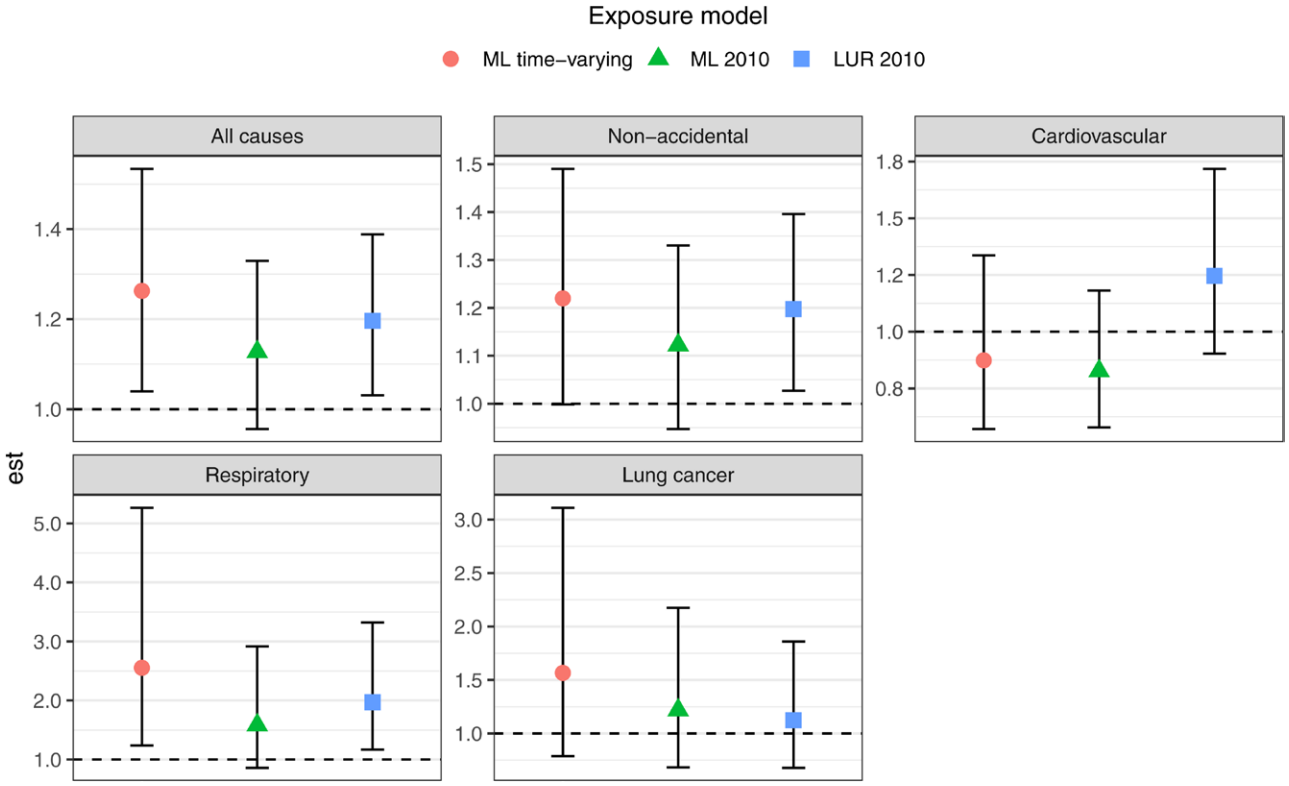

FIGURE 3. Comparison of hazard ratios (HRs, with 95% confidence intervals) of different mortality causes associated with an increase of 10 μg/m³ in $PM_{2.5}$, estimated using alternative exposure sources and indices, using the fully adjusted confounder model (Model 6 in Table 3).

Kingdom.[21–23] A potential explanation is the relatively young age of the UKB participants, given the specific recruitment constraints (40–69 years of age) and the still relatively short follow-up, although this issue should be studied further in future analyses. In general, for all outcomes, differences can be due to variations in population susceptibility and specific PM composition. It must also be noted that, in any case, the relatively wide confidence intervals include ranges more consistent with the literature.

Important strengths of this study are the fine control for confounding and the use of time-varying exposure information. Most of the studies in the literature are based on large administrative cohorts with limited individual-level information,[13,14,17] or the pooling of small and heterogeneous cohorts,[16,22] although with some exceptions.[34] The UKB database offers a unique combination where information on both contextual- and individual-level characteristics was collected for a large sample of half a million participants. The results reported in Table 3 demonstrate the potential biases arising from incomplete confounding control, with area-level deprivation, rural–urban settings, and green space, as well as individual lifestyle factors, all featuring as important predictors and potential confounders. Similarly, the analysis takes advantage of time-varying exposure information, provided by the detailed reconstruction of individual-level daily exposure histories to $PM_{2.5}$ for the full cohort using state-of-the-art spatiotemporal ML models.[27] With some exceptions,[14,35] most of the published studies used simpler effect summaries, based for instance on time-invariant exposure indices (for instance, the average $PM_{2.5}$ in a given year),[16,17] or alternatively using time-varying measures represented only by the current year's level.[13] Previous methodologic works highlighted the benefit of using time-varying exposure summaries based on longer exposure windows and reconstructed using exposure models with high spatial and temporal resolution.[20,36] This evidence is consistent with the results of our sensitivity analysis in Figure 3.

An original aspect of this study is the application of flexible methods to assess the temporal structure of the association and identify relevant exposure windows across different outcomes. This issue has been addressed in the large epidemiologic literature for other risk factors such as smoking,[37] radon,[38] or medications,[39] but apart from some exceptions,[40,41] it has been notably overlooked in the research on the long-term risks of air pollution. In particular, the application of DLMs offers an easily adaptable framework where different functions can be applied to model the lag–response relationship, and it takes full advantage of finely temporally

© 2024 Wolters Kluwer Health, Inc. All rights reserved.

Downloaded from http://journals.lww.com/epidem by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8K2+Ya6H515kE= on 11/26/2024

Copyright © 2024 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

disaggregated exposure histories. The distributed lag model methodology was originally implemented and is commonly used in time series studies to analyze short-term risks of environmental factors.[29,42,43] More recently, it has been generalized beyond time series data and it is now applicable for the analysis of long-term associations,[30] although with limited applications so far.[41] Our study demonstrates its potential, and although the large uncertainty in the results prevents drawing firm conclusions in this specific case, it can be replicated in future follow-up updates of the UKB cohort and other studies.

An important limitation of the UKB cohort is its voluntary participation with a low response rate of 5.5%,[44] resulting in a selection of participants nonrepresentative of the UK population.[45] While it has been established that representativeness is not a requirement for valid epidemiologic inference,[46,47] a strong selection can result in biased estimates in the presence of unmeasured risk factors that act as colliders when associated with both the exposure and the probability of selection.[31,47,48] While we cannot verify this hypothesis, we included a washout period in a sensitivity analysis, recommended in a recent study as a measure to attenuate the bias.[31] This resulted in higher effect estimates, indicating that in case this bias occurred, our estimates would be conservative.

Other limitations must be acknowledged. First, despite the size of the UKB cohort, its follow-up is still relatively short and results in a small number of health events compared to other published studies. This affected the precision of the estimates, in particular for cause-specific mortality outcomes, and prevented supplementary analysis such as on the shape of the exposure–response curve and potential nonlinearities. In addition, most of the potential confounders included in the analysis were only measured at baseline, and we could not account for potential changes during the follow-up, in particular, related to lifestyle factors. Regarding the exposure measurement, while we relied on high-resolution $PM_{2.5}$ levels derived from the spatiotemporal ML model, these only represent residential concentrations, which can result in some degree of exposure misclassification. Furthermore, the exposure reconstruction could only cover the period from 2003 on, thus limiting the temporal windows we could assess in our analysis. Finally, we were not able to adjust for copollutants such as nitrogen dioxide (NO_2) or ozone (O_3). However, the linkage of these and other environmental stressors to the UKB database is ongoing and this limitation can be addressed in future analyses.

In conclusion, this epidemiologic assessment examined the association between long-term exposure to $PM_{2.5}$ and mortality in the UKB cohort, a large prospective study of more than half a million British adults, using finely reconstructed exposures, strict control for confounding, and advanced study designs and statistical methods. The results indicate substantial risks associated with $PM_{2.5}$ for all-cause, nonaccidental, and respiratory causes, as well as lung cancer, but not for cardiovascular outcomes.

Downloaded from http://journals.lww.com/epidem by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8K2+Ya6H515kE= on 11/26/2024

ACKNOWLEDGMENTS

This research has been conducted using the UK Biobank Resource under application number 56431. This work was part of a PhD project supported by the Nagasaki University WISE Programme.

REFERENCES

1. Burnett R, Chen H, Szyszkowicz M, et al. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proc Natl Acad Sci USA*. 2018;115:9592–9597.
2. Wu X, Braun D, Schwartz J, Kioumourtzoglou MA, Dominici F. Evaluating the impact of long-term exposure to fine particulate matter on mortality among the elderly. *Sci Adv*. 2020;6:eaba5692.
3. Hoffmann B, Brunekreef B, Andersen ZJ, Forastiere F, Boogaard H. Benefits of future clean air policies in Europe. *Environ Epidemiol*. 2022;6:e221.
4. Rajagopalan S, Al-Kindi SG, Brook RD. Air pollution and cardiovascular disease: JACC state-of-the-art review. *J Am Coll Cardiol*. 2018;72:2054–2070.
5. Doiron D, de Hoogh K, Probst-Hensch N, et al. Air pollution, lung function and COPD: results from the population-based UK Biobank study. *Eur Respir J*. 2019;54:1802140.
6. Hamra GB, Guha N, Cohen A, et al. Outdoor particulate matter exposure and lung cancer: a systematic review and meta-analysis. *Environ Health Perspect*. 2014;122:906–911.
7. Afsar B, Afsar RE, Kanbay A, Covic A, Ortiz A, Kanbay M. Air pollution and kidney disease: review of current evidence. *Clin Kidney J*. 2018;12:19–32.
8. Fu P, Guo X, Cheung FMH, Yung KKL. The association between $PM_{2.5}$ exposure and neurological disorders: a systematic review and meta-analysis. *Sci Total Environ*. 2019;655:1240–1248.
9. Yang B-Y, Fan S, Thiering E, et al. Ambient air pollution and diabetes: a systematic review and meta-analysis. *Environ Res*. 2020;180:108817.
10. Chen J, Hoek G. Long-term exposure to PM and all-cause and cause-specific mortality: a systematic review and meta-analysis. *Environ Int*. 2020;143:105974.
11. Pope CA, Coleman N, Pond ZA, Burnett RT. Fine particulate air pollution and human mortality: 25+ years of cohort studies. *Environ Res*. 2020;183:108924.
12. Krewski D, Burnett R, Goldberg M, et al. Overview of the reanalysis of the Harvard six cities study and American cancer society study of particulate air pollution and mortality. *J Toxicol Environ Health A*. 2003;66:1507–1552.
13. Di Q, Wang Y, Zanobetti A, et al. Air pollution and mortality in the Medicare population. *N Engl J Med*. 2017;376:2513–2522.
14. Crouse DL, Peters PA, Hystad P, et al. Ambient $PM_{2.5}$, O_3, and NO_2 exposures and associations with mortality over 16 years of follow-up in the Canadian Census Health and Environment Cohort (CanCHEC). *Environ Health Perspect*. 2015;123:1180–1186.
15. Beelen R, Raaschou-Nielsen O, Stafoggia M, et al. Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. *Lancet*. 2014;383:785–795.
16. Strak M, Weinmayr G, Rodopoulou S, et al. Long term exposure to low level air pollution and mortality in eight European cohorts within the ELAPSE project: pooled analysis. *BMJ*. 2021;374:n1904.
17. Stafoggia M, Oftedal B, Chen J, et al. Long-term exposure to low ambient air pollution concentrations and mortality among 28 million people: results from seven large European cohorts within the ELAPSE project. *Lancet Planet Health*. 2022;6:e9–e18.
18. de Hoogh K, Chen J, Gulliver J, et al. Spatial $PM_{2.5}$, NO_2, O_3 and BC models for Western Europe - evaluation of spatiotemporal stability. *Environ Int*. 2018;120:81–92.
19. van Donkelaar A, Martin RV, Li C, Burnett RT. Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. *Environ Sci Technol*. 2019;53:2595–2611.
20. Crouse DL, Erickson AC, Christidis T, et al. Evaluating the sensitivity of $PM_{2.5}$-mortality associations to the spatial and temporal scale of exposure assessment. *Epidemiology*. 2020;31:168–176.

© 2024 Wolters Kluwer Health, Inc. All rights reserved.

Copyright © 2024 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

21. Carey IM, Atkinson RW, Kent AJ, van Staa T, Cook DG, Anderson HR. Mortality associations with long-term exposure to outdoor air pollution in a national English cohort. *Am J Respir Crit Care Med*. 2013;187:1226–1233.
22. Dehbi H-M, Blangiardo M, Gulliver J, et al. Air pollution and cardiovascular mortality with over 25years follow-up: a combined analysis of two British cohorts. *Environ Int*. 2017;99:275–281.
23. Wang M, Zhou T, Song Q, et al. Ambient air pollution, healthy diet and vegetable intakes, and mortality: a prospective UK Biobank study. *Int J Epidemiol*. 2022;51:1243–1253.
24. Sudlow C, Gallacher J, Allen N, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015;12:e1001779.
25. Conroy M, Sellors J, Effingham M, et al. The advantages of UK Biobank's open-access strategy for health research. *J Intern Med*. 2019;286:389–397.
26. Eeftens M, Beelen R, de Hoogh K, et al. Development of land use regression models for PM(2.5), PM(2.5) absorbance, PM(10) and PM(coarse) in 20 European study areas; results of the ESCAPE project. *Environ Sci Technol*. 2012;46:11195–11205.
27. Libardi AC, Masselot P, Schneider R, et al. High resolution mapping of nitrogen dioxide and particulate matter in Great Britain (2003-2021) with multi-stage data reconstruction and ensemble machine learning methods. *Atmos Pollut Res*. 2024;15:102284.
28. Vanoli J, Mistry MN, De La Cruz Libardi A, et al. Reconstructing individual-level exposures in cohort analyses of environmental risks: an example with the UK Biobank [published online ahead of print]. *J Expo Sci Environ Epidemiol*. 2024. doi:10.1038/s41370-023-00635-w.
29. Gasparrini A, Armstrong B, Kenward MG. Distributed lag non-linear models. *Stat Med*. 2010;29:2224–2234.
30. Gasparrini A. Modeling exposure-lag–response associations with distributed lag non-linear models. *Stat Med*. 2014;33:881–899.
31. Chen C, Chen H, Kaufman JS, Benmarhnia T. Differential participation, a potential cause of spurious associations in observational cohorts in environmental epidemiology. *Epidemiology*. 2024;35:174–184.
32. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *BMC Med Res Methodol*. 2009;9:57.
33. Pappin AJ, Christidis T, Pinault LL, et al. Examining the shape of the association between low levels of fine particulate matter and mortality across three cycles of the Canadian Census Health and Environment Cohort. *Environ Health Perspect*. 2019;127:107008.
34. Pinault L, Tjepkema M, Crouse DL, et al. Risk estimates of mortality attributed to low concentrations of ambient fine particulate matter in the Canadian community health survey cohort. *Environ Health*. 2016;15:18.
35. Chen J, Braun D, Christidis T, et al. Long-term exposure to low-level $PM_{2.5}$ and mortality: investigation of heterogeneity by harmonizing analyses in large cohort studies in Canada, United States, and Europe. *Environ Health Perspect*. 2023;131:127003.
36. Lefler JS, Higbee JD, Burnett RT, et al. Air pollution and mortality in a large, representative U.S. cohort: multiple-pollutant analyses, and spatial and temporal decompositions. *Environ Health*. 2019;18:101.
37. Richardson DB, Terschuren C, Pohlabeln H, Jockel KH, Hoffmann W. Temporal patterns of association between cigarette smoking and leukemia risk. *Cancer Causes Control*. 2008;19:43–50.
38. Langholz B, Thomas D, Xiang A, Stram D. Latency analysis in epidemiologic studies of occupational exposures: application to the Colorado Plateau uranium miners cohort. *Am J Ind Med*. 1999;35:246–256.
39. Abrahamowicz M, Bartlett G, Tamblyn R, du Berger R. Modeling cumulative dose and exposure duration provided insights regarding the associations between benzodiazepines and injuries. *J Clin Epidemiol*. 2006;59:393–403.
40. Schwartz J, Coull B, Laden F, Ryan L. The effect of dose and timing of dose on the association between airborne particles and survival. *Environ Health Perspect*. 2008;116:64–69.
41. Kriit HK, Andersson EM, Carlsen HK, et al. Using distributed lag non-linear models to estimate exposure lag–response associations between long-term air pollution exposure and incidence of cardiovascular disease. *Int J Environ Res Public Health*. 2022;19:2630.
42. Schwartz J. The distributed lag between air pollution and daily deaths. *Epidemiology*. 2000;11:320–326.
43. Armstrong B. Models for the relationship between ambient temperature and daily mortality. *Epidemiology*. 2006;17:624–631.
44. Swanson JM. The UK Biobank and selection bias. *Lancet*. 2012;380:110.
45. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am J Epidemiol*. 2017;186:1026–1034.
46. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol*. 2013;42:1012–1014.
47. Richiardi L, Pizzi C, Pearce N. Commentary: representativeness is usually not necessary and often should be avoided. *Int J Epidemiol*. 2013;42:1018–1022.
48. Pizzi C, De Stavola B, Merletti F, et al. Sample selection and validity of exposure-disease association estimates in cohort studies. *J Epidemiol Commun Health*. 2010;65:407–411.

Downloaded from http://journals.lww.com/epidem by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8K2+Ya6H515kE= on 11/26/2024

© 2024 Wolters Kluwer Health, Inc. All rights reserved.

Copyright © 2024 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.