International Journal of Epidemiology, 2026, **55(1)**, dyaf214
https://doi.org/10.1093/ije/dyaf214
Original Article

IEA
International Epidemiological Association
OXFORD

Original Article

Exposure measurement error in air-pollution epidemiology and its determinants: results from the MELONS study

Dimitris Evangelopoulos[1,2,*], Dylan Wood[1,2], Ben Barratt[1,2], Hanbin Zhang[1,3], Audrey de Nazelle[4], Sean Beevers[1,2], Barbara K. Butland[5], Evangelia Samoli[6], Joel Schwartz[7,8], Kees de Hoogh[9,10], Konstantina Dimakopoulou[6,11], Heather Walton[1,2], Klea Katsouyanni[1,2,6]

[1]Environmental Research Group, MRC Centre for Environment and Health, Faculty of Medicine, Imperial College London, London, United Kingdom
[2]NIHR HPRU in Environmental Exposures and Health, Imperial College London, London, United Kingdom
[3]European Centre for Environment and Human Health, University of Exeter, Exeter, United Kingdom
[4]Centre for Environmental Policy, Imperial College London, London, United Kingdom
[5]School of Health & Medical Sciences, City St George's, University of London, London, United Kingdom
[6]Department of Hygiene, Epidemiology and Medical Statistics, National and Kapodistrian University of Athens, Medical School, Athens, Greece
[7]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, United States
[8]Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, United States
[9]Swiss Tropical and Public Health Institute, Allschwil, Switzerland
[10]University of Basel, Basel, Switzerland
[11]Department of Basic and Clinical Sciences, University of Nicosia Medical School, Nicosia, Cyprus

*Corresponding author: Environmental Research Group, MRC Centre for Environment and Health, Faculty of Medicine, Imperial College London, White City Campus, Wood Lane, London W12 0BZ, United Kingdom. E-mail: d.evangelopoulos@imperial.ac.uk

Abstract

Introduction: In air-pollution epidemiology, measured or modelled surrogate exposure estimates, prone to measurement error (ME), are used to investigate the health effects of exposure to pollution of outdoor origin, potentially leading to biased effect estimates. We predicted the annual personal exposure from outdoor sources by using personal measurements, compared it with concentrations from surrogate metrics, and quantified the ME magnitude, type, and determinants.

Methods: We used measurements from four panel studies in London, UK, and predicted personal exposures to fine particulate matter ($PM_{2.5}$), nitrogen dioxide (NO_2), ozone (O_3), and black carbon (BC). We compared those with surrogate exposures, including measurements from fixed-site monitors, modelled ambient concentrations, or hybrid methods accounting for people's mobility. We estimated the exposure ME magnitude, correlations, and variance ratios between surrogate measures and personal exposure, and the percentages of classical/Berkson-type errors. Individual- and area-level characteristics, such as age, sex, socio-economic status, and time spent outdoors, were assessed as potential error determinants.

Results: Predicted annual personal exposures to $PM_{2.5}$, NO_2, O_3, and BC from outdoor sources were overestimated by surrogate metrics, with mean differences of up to 10.1, 40.0, 61.7, and 2.6 $\mu g/m^3$, respectively. The variance ratios and Pearson correlation coefficients between surrogate and predicted personal exposures ranged from 0.03 to 165.02 and –0.24 to 0.25. Time–activity adjustment reduced errors substantially. Berkson-type errors dominated the ME for $PM_{2.5}$ and BC (43%–81% and 26%–98%, respectively), whilst classical errors characterized gases (>94% for both NO_2 and O_3). Time spent outdoors, house type, and deprivation were associated with exposure error.

Conclusion: The use of surrogate exposures to investigate the health effects of long-term exposure to air pollution from outdoor sources may bias the epidemiological estimates due to ME. Information about the error structures and their determinants can be used for correction and the identification of the true exposure–response functions.

Keywords: air pollution; measurement error; mixture error; error determinants

Received: 12 June 2025; **Accepted:** 30 October 2025

Key Messages

- This study assessed exposure measurement error (ME) by comparing ambient measurements or modelled concentrations with long-term personal exposures to outdoor air pollution.
- ME was large in magnitude and variability, correlations between personal and surrogate exposures were low, and gases were characterized by classical errors while particles were characterized by Berkson errors.
- Exposure ME can lead to biased health-effect estimates and these findings can inform future research that aims to account for ME in study designs and epidemiological analyses.

Introduction

Over the last decades, epidemiological cohort studies have examined the associations between multiple health outcomes and long-term exposure to air pollution from outdoor sources [1, 2]. However, these studies rely on surrogate exposure assessment methods, such as measurements from fixed-site monitors or modelled ambient concentrations, due to the inherent challenges in obtaining precise long-term personal exposure data for large populations [3]. While these surrogate exposures serve as invaluable tools in identifying associations with health outcomes and informing policy-making on the importance of ambient air-pollution reduction measures, they are prone to exposure error, as they cannot entirely capture individual exposures. The difference between a 'gold-standard' measure, such as personal exposure to pollutants from outdoor sources, and surrogate measures can be influenced by factors such as occupation, time–activity, the infiltration efficiency of buildings, and socio-economic status [4]. For personal exposures from outdoor sources specifically, it must be noted that measurements are extremely difficult—almost impossible—to implement.

A limited number of studies have attempted to investigate the differences between personal and surrogate exposures, in terms of both their differences in scale, as well as their correlation [5, 6]. Even fewer studies have examined these relationships by using personal exposure from solely outdoor sources as the 'gold-standard' exposure, due to the challenges inherent in its estimation [4, 7]. Discrepancies between personal and surrogate exposures are generally evaluated on a short-term scale, using daily data for short periods of time. More importantly, no study to our knowledge has provided a holistic assessment of the exposure measurement error (ME) structures, including information on the error magnitude, variability, and type for different pollutants and exposure assessment methods, the correlation and variance ratios between surrogate and true exposures, as well as the correlations between the errors of different air pollutants. The type (classical or Berkson) of error depends on the pollutant, its propagation and sources, and how it infiltrates buildings, and can also vary across each error-prone exposure assessment method. Particles and gases are different in all these aspects, so their error structures are also expected to be different.

Exposure error is also important for epidemiological analysis, as the error magnitude, variability, and type are the main drivers of exposure ME bias in health-effect estimation [8, 9]. Simulation studies have examined multiple scenarios for a range of these driving factors, including varying the percentages of classical and Berkson errors in the mixture. The biases reported were generally to the null and of different magnitudes, ranging from <1% to >50%, depending on what was defined as 'gold-standard' exposure and the error parameters [8, 10]. However, only a few studies have assessed ME bias by using personal exposure from outdoor sources as the 'gold-standard' exposure, and only for short-term exposures [11, 12]. Estimating the impacts of personal exposures from outdoor sources is important to inform national and international air-pollution guidelines, policies, and legislation [13].

The hypotheses tested in previous theoretical and simulation studies are mainly for classical-type errors for fine particulate matter ($PM_{2.5}$) and nitrogen dioxide (NO_2). Moreover, the inputs used in the simulation analyses were informed by limited comparisons between modelled concentrations (surrogate) and measurements from fixed-site monitors ('gold-standard') or by arbitrary values for the drivers of exposure ME bias in epidemiological analyses [9, 10]. The number of classical and Berkson components in the error mixture, as well as an approximation of the correlation and variance ratios between surrogate and personal exposures or correlations between exposure errors based on real data, has not been investigated. Other pollutants, such as ozone (O_3) and black carbon (BC), have also not been assessed previously. In addition, previous reviews have shown that a very limited proportion of air-pollution epidemiological studies perform ME bias corrections [14, 15]. Accurate estimates of the error parameters for multiple air pollutants can inform future studies aiming to quantify biases in epidemiological associations due to ME and lead to improved correction methods.

To bridge these gaps in the literature, we estimated data on personal exposure from outdoor sources (i.e. hypothesized 'gold-standard' exposure) from four exposure panel studies in London, UK [16–19] and predicted the annual exposures for particles and gaseous pollutants by using an extrapolation analysis. We explored the disparities between surrogate metrics and personal exposures to outdoor air pollutants. Our surrogates included correcting for individual time–activity patterns for the same age group and area of residence based on a survey—something rarely done in epidemiological cohort studies. In addition, we investigated characteristics that determine the magnitude of exposure error and decomposed the error into classical and Berkson components by estimating their contributions in the error mixture. The overall aim was to gain insight into the error structures of $PM_{2.5}$, NO_2, O_3, and BC, and multiple surrogate exposures generally used in air-pollution studies. Our analysis provides inputs on the error magnitude and variability when surrogate exposures are used in epidemiological analyses that aim to account and correct for exposure ME bias, but may not have access to 'gold-standard' exposure assessment methods from validation sub-studies.

Methods

Personal exposure measurement campaigns

We used data from four previous campaigns in London of diverse demographics, with total personal exposure measurements for air pollution and GPS data. More specifically, we utilized measurements for 10 950 person-days of 75 chronic obstructive pulmonary disease (COPD) patients aged 50–90 years ['Characterisation of COPD exacerbations using environmental exposure modelling' (COPE) study; June 2015 to October 2017] [16], 824 person-days of 38 healthy adults aged 18–60 years ['Physical Activity through Sustainable Transport Approaches' (PASTA) study; April 2015 to March 2016] [17], 820 person-days of 162 primary schoolchildren and teachers ['Breathe London Wearables' (BLW) study; March to May 2019] [18], and 307 person-days of 66 professional drivers ['The Diesel Exposure Mitigation Study' (DEMiST) study; February 2018 to January 2019] [19]. In COPE, $PM_{2.5}$, NO_2, and O_3 measurements were collected for an average of 144 days per participant. In BLW, $PM_{2.5}$ personal exposures were collected for five consecutive weekdays by using backpacks with built-in air-quality sensors. In PASTA and DEMiST, BC measurements were collected for 5 consecutive days in three seasons (i.e. 15 days in total) and an average of 4.5 consecutive workdays, respectively.

All participants were provided with a personal air-quality monitor that also recorded GPS positions. We classified their location every 15 minutes into microenvironments including 'home', 'other indoor', and 'outdoor/transit'. Personal air-pollution exposures were also collected at 15-minute resolution and averaged at 1-hour resolution, while the location assigned was the one in which the participant spent most of the 1-hour time period. All measurement campaigns have been through proper quality assurance, quality control (QA/QC) processes and the sensors and methodology used are described elsewhere [18–21]. In general, the agreement between sensors and reference monitors (expressed as R^2 values) was good and comparable across all four measurement campaigns. For COPE, the R^2 value was >0.8 for NO_2 and O_3, as well as 0.8 during summer and 0.6 during winter for $PM_{2.5}$. For BLW ($PM_{2.5}$ only), DEMiST and PASTA (BC only), the R^2 values were >0.7.

Estimation of personal exposure from outdoor sources during the measurement period

We further collected ambient pollution data at 1-hour resolution by using measurements from the London Air Quality Network (www.londonair.org.uk) and estimates from the London Air Quality Toolkit model [22]. Personal exposure measurements when participants were at home (school for BLW and work for DEMiST) and matched ambient data were used to estimate the participant-, pollutant-, and month-specific infiltration efficiency. Only indoor data points when no indoor source event occurred (identified by following the methods described in a previous publication [23]) were included in the infiltration efficiency calculation. The COPE dataset provided data across the whole year; thus, we assessed the monthly variation of infiltration efficiencies (Supplementary Fig. A1).

Our 'gold-standard' exposure to ambient pollution was the personal exposure to pollution from outdoor sources. To estimate it during the measurement period, we calculated the proportion of ambient levels infiltrating into the home by using the infiltration efficiencies described above when participants were indoors. When they were outdoors or in other indoor environments, i.e. 'not at home', assuming no indoor sources, personal exposure from outdoor sources was equal to the measured personal exposure. The exposure separation methodology during the measurement period is described in Zhang *et al.* (2025) [23].

Predicted annual personal exposures from outdoor sources

For the prediction of annual personal exposures from outdoor sources, we combined time–activity patterns and infiltration efficiencies during the measurement period with ambient pollution measurements for a full year from the campaign start date for each participant. We predicted whether a person was likely to be 'at home' or 'not at home' during the extrapolation period by utilizing a random forest behaviour prediction model, using the day of the week and hour of the day as inputs, and a Synthetic Minority Over-sampling Technique that accounted for imbalanced distribution of the measurement-period data (e.g. >90% 'at home' for COPE) [24]. For BLW, we adjusted the time–activity of the schoolchildren based on child time use surveys [25, 26]. More details on the annual extrapolation are provided in Supplementary Section A. The methods applied were consistent across all pollutants included in the present analysis, i.e. $PM_{2.5}$, NO_2, O_3, and BC, but the infiltration efficiencies and ambient levels used were pollutant-specific.

Derivation of surrogate exposures

We estimated surrogate exposures (W_i) to $PM_{2.5}$, NO_2, O_3, and BC for the four campaigns' participants, including predictions from spatio-temporal models, measurements from the nearest fixed-site monitor, and hybrid methods that accounted for people's mobility. More specifically, we collected urban background and roadside site measurements for $PM_{2.5}$, NO_2, O_3, and BC from 14, 42, 12, and 2 London Air Quality Network monitors, respectively, within the Greater London area, using the '*openair*' R package [27]. For each individual, we identified the nearest monitor to their residence for each pollutant under investigation and estimated long-term exposures by averaging the hourly concentrations (i) for 1 year from the date on which each individual joined the personal monitoring campaign and (ii) for the measurement period of each campaign. We excluded kerbside and industrial sites.

For modelled concentrations of $PM_{2.5}$, NO_2, and O_3, we used the 'Comparative evaluation of Spatio-Temporal Exposure Assessment Methods for estimating health effects of air pollution' STEAM model 2009–13 estimates [28], which utilized a generalized additive model that combined data from spatio-temporal land-use regression and the Community Multiscale Air Quality urban (CMAQ-urban [29]) chemical transport model and, additionally for $PM_{2.5}$ only, a prediction model using satellite data on aerosol optical depth [30, 31]. Models provided daily predictions, which were averaged annually and then for 2009–13 to assess long-term exposures at the postcode centroid level (the median number of households per postcode in England and Wales is 14 [32]) and assigned to the participants based on their postcode of residence. The 10-fold cross-validation R^2 values were 0.80 for NO_2, 0.79 for $PM_{2.5}$, and 0.75 for O_3 [28]. For BC, we used the 'Effects of Low-Level Air Pollution: A Study in Europe' (ELAPSE) model, which incorporated annual mean BC concentrations from monitors, $PM_{2.5}$ satellite data, and chemical transport model data [33]. Greater London-wide 2010 BC estimates were available at a resolution of 100×100 m and were assigned to individuals based on their residential address. The model had an overall R^2 value of 0.54. In addition, to account for the temporal misalignment between the original modelled estimates, i.e. 2009–13 for STEAM and 2010 for ELAPSE, and the personal exposure campaigns (2015–19), and the downward trend of air pollution across the years, the modelled estimates were adjusted for each individual by the ratio of the nearest-monitor measurements during 2015–19 over those of 2009–13 ($PM_{2.5}$, NO_2, and O_3 for STEAM) or 2010 (BC for ELAPSE), as done previously in epidemiological studies [34], and these are referred to as temporally aligned STEAM/ELAPSE estimates.

We created another surrogate exposure by considering typical time–activity patterns by age group and area of residence. Specifically, we used the London Hybrid Exposure Model (LHEM), which combines time–activity for Londoners with CMAQ-urban (2011)-modelled ambient $PM_{2.5}$ and NO_2 concentrations, in-building- and in-vehicle-modelled concentrations, and estimates of personal exposure to pollution from outdoor sources reflective of time–activity and microenvironment exposure [35]. The LHEM adjustment was applied to the nearest-monitor measurements and the STEAM-modelled concentrations described above. More details are provided in Supplementary Section A.

The above surrogate exposures were assigned to the study participants and a harmonized database of personal and surrogate exposures was constructed.

Statistical analysis

Exposure ME was defined as the difference between the long-term averages of each surrogate, W_i, and the predicted annual personal exposure from outdoor sources, X, i.e. $W_i - X$. We estimated the error magnitude from the mean and median values, and error variability from the standard deviation and interquartile range (IQR). We also calculated the Spearman correlation coefficients between surrogate and personal exposures to account for potential non-linearities. Scatter plots were created to assess the shape of the relationships and whether the observed errors were heteroscedastic. We also used estimated personal exposure from outdoor sources during the measurement period of the campaigns or when individuals were only indoors, to compare with predicted annual personal exposures. The former assessed the impact of the annual extrapolation of the personal exposures, while the latter indicated whether individuals' mobility increased the discrepancies between personal exposures and the static surrogates at the participants' residence. Finally, to assess the impact of the exposure separation into indoor- and outdoor-generated pollution, we used the total personal exposure during the measurement period of each campaign (not extrapolated annually) and compared that with the surrogates described above.

We also calculated the covariance, Pearson correlation coefficient, and variance ratio between W_i and X. We further estimated the Pearson correlation coefficients of the exposure errors for $PM_{2.5}$ and NO_2, $PM_{2.5}$ and O_3, and NO_2 and O_3 (COPE only), as the correlations between the errors are driving factors of the ME bias in multi-pollutant models [9]. We used the latent variable mixed error model described by Carroll *et al.* (2006) [36] and estimated the percentages of classical and Berkson components in the error mixture, by pollutant, measurement campaign, and surrogate measure. We also conducted variogram analysis and calculated all the components of a typical semi-variogram (i.e. nugget, sill, and practical range) by using exponential, spherical, and Matérn models. For these models, we calculated the nugget by squaring the standard error of the mean of the daily personal exposures from outdoor sources, which reflects the random error in the estimation of long-term, annual averages from daily data and can be attributed to within-person variation.

To identify potential exposure error determinants, we used individual and shared characteristic data measured for COPE (NO_2, $PM_{2.5}$, and O_3) and PASTA (BC), as BLW and DEMiST did not provide information on any potential error determinants. More specifically, we applied linear regression models, using the exposure ME (in $\mu g/m^3$) of each surrogate and each pollutant as the dependent variable. A list of factors that may explain the error variability (collected in the measurement campaigns) was assessed, including age (in years) from baseline questionnaires, Index of Multiple Deprivation (IMD) 2019 [37] (unitless), and measured daily average time spent outside the home (in hours), as well as gender (female/male), residence type (flat/apartment, detached/semi-detached and other, including bungalows/cottages/terraced houses/mobile homes), and cooking fuel (gas, electric and other, including a mix of gas hob and electric oven or wood burning) reported in the baseline questionnaires. Additionally for COPE, a categorical variable denoting the doctor-diagnosed COPD status severity was adjusted for (mild/moderate, severe, and very severe). For PASTA, the self-reported daily percentage of time spent physically active and the body mass index (BMI) (in kg/m^2) from baseline questionnaires were adjusted for as continuous variables, as well as smoking status (never smoker and former smoker; no current smokers were recruited) and self-reported worry regarding air pollution ('not worried' and 'worried'). These or other potential error determinants were not available in BLW and DEMiST.

Predicted personal and surrogate exposure measures were standardized by subtracting the mean and dividing by the standard deviation, in order to compare between them at the same scale before calculating the exposure errors ($W_{istd} - X_{std}$). Standardization was applied because ambient concentrations are considerably higher than the personal exposure from outdoor sources, i.e. a systematic difference, and the variance was also larger. Observations with missing values were excluded from the analysis. Outliers in the error estimation that could have impacted the regression estimates were identified as values above $Q3 + 3\times IQR$ (where Q3 is the 75th percentile) and removed. The total explained variability of the models was assessed with the adjusted R^2 statistic. The analysis was conducted by using R, version 4.2.1.

Results

Description of 'gold-standard' personal and surrogate exposures

The mean (standard deviation) predicted annual personal exposure from outdoor sources was 4.5 (1.5) μg/m^3 for NO_2, 5.6 (2.0) and 5.0 (1.7) μg/m^3 for $PM_{2.5}$ for COPE and BLW, respectively, 2.9 (1.1) μg/m^3 for O_3, and 1.4 (1.0) and 1.3 (0.7) μg/m^3 for BC (PASTA and DEMiST, respectively). The corresponding values for the measurement-period personal exposures were very similar, while the surrogate measures

Table 1. Means and standard deviations of personal exposure from outdoor sources (annually predicted and during the measurement period) and assigned surrogate concentration estimates in four exposure measurement campaigns (COPE, BLW, PASTA, and DEMiST).

Mean concentration estimate ± standard deviation (μg/m^3)						
Exposure estimate	**NO_2**	**$PM_{2.5}$**		**O_3**	**BC**	
	COPE	**COPE**	**BLW**	**COPE**	**PASTA**	**DEMiST**
			Personal exposure estimate			
Personal from outdoors (annually predicted)	4.5 ± 1.5	5.6 ± 2.0	5.0 ± 1.7	2.9 ± 1.1	1.4 ± 1.0	1.3 ± 0.7
Personal from outdoors (measurement period)	4.0 ± 1.3	5.1 ± 3.0	5.5 ± 4.3	2.7 ± 1.1	1.2 ± 0.5	1.7 ± 1.0
			Surrogate exposure estimate			
Nearest monitor (annual)	44.5 ± 18.1	12.6 ± 2.3	10.0 ± 1.7	32.6 ± 7.9	4.0 ± 1.7	1.8 ± 0.8
STEAM (2009–13) for NO_2, $PM_{2.5}$, and O_3/ELAPSE (2010) for BC	40.1 ± 18.7	15.6 ± 1.4	16.3 ± 1.4	60.0 ± 7.2	2.2 ± 0.3	2.1 ± 0.3
Temporally aligned STEAM/ELAPSE (2015–19)	33.2 ± 14.8	11.3 ± 1.1	11.7 ± 1.1	64.6 ± 6.5	1.0 ± 0.2	1.0 ± 0.2
LHEM-adjusted STEAM (2009–13)	13.5 ± 4.3	9.4 ± 1.0	10.0 ± 0.9	Not applicable	Not applicable	Not applicable

Measurement period and number of participants (n): COPE: June 2015 to October 2017, $n = 75$. BLW: March to May 2019, $n = 162$. PASTA: April 2015 to March 2016, $n = 38$. DEMiST: February 2018 to January 2019, $n = 66$.

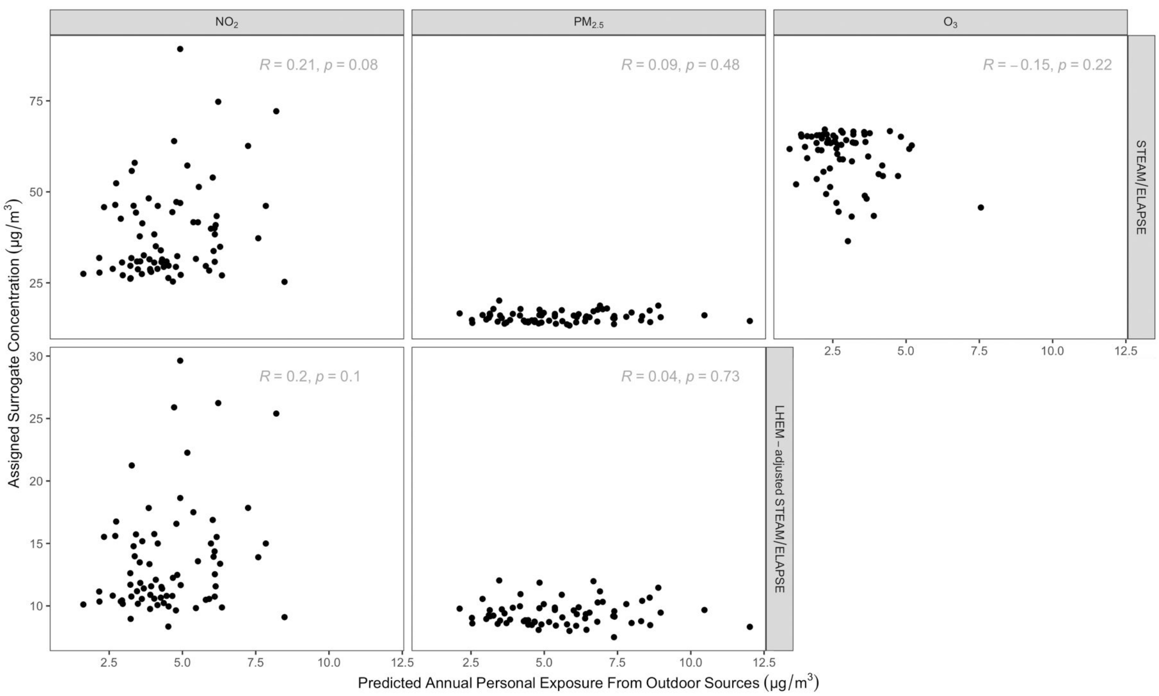

Figure 1. Scatter plots of predicted annual personal exposure from outdoor sources, STEAM-modelled ambient concentrations, and the corresponding LHEM-adjusted exposure estimates for COPE participants for NO_2, $PM_{2.5}$, and O_3. R represents the Spearman correlation coefficient; p represents the corresponding correlation P value.

assessing residential outdoor concentrations were higher in magnitude (Table 1). The LHEM-adjusted surrogate accounted for the time–activity and indoor concentrations, and thus was more comparable but larger than the predicted annual personal exposures. Correspondingly, standard deviations of the surrogate measures were generally larger than those of personal exposures for NO_2 and O_3, but not for $PM_{2.5}$ and BC. These differences in the scale (magnitude) and variability of measurements are illustrated in Fig. 1 (and Supplementary Fig. A4) for STEAM-modelled concentrations (NO_2, $PM_{2.5}$, and O_3) and LHEM-adjusted-modelled estimates (NO_2 and $PM_{2.5}$) assigned to COPE study participants. Bivariate relationships for the predicted annual or measured personal exposures and all the surrogates collected for the present study are provided as scatter plots in Supplementary Figs B2 and B3 (NO_2), B7 and B8 ($PM_{2.5}$), B12 and B13 (O_3), and B17 and B18 (BC).

Assessment of ME

The Spearman correlation coefficients between the surrogate and predicted annual personal exposures from outdoor sources by pollutant and study were generally low, with the highest correlations observed for the COPE and NO_2-modelled estimates ($r = 0.21$) (Table 2). Low correlations were also observed for the measurement period only, i.e. without extrapolation (Table 2), due to the good agreement between the predicted annual and measurement-period personal exposure estimates, especially for our longer-exposure campaign, i.e. COPE

Table 2. Spearman correlation coefficients between personal exposure from outdoor sources (for the measurement period and annually predicted) and assigned surrogate concentration estimates in four exposure measurement campaigns (COPE, BLW, PASTA, and DEMiST).

Surrogate exposure	NO2	PM2.5		O3	BC	
	COPE	COPE	BLW	COPE	PASTA	DEMiST
Spearman correlation coefficient for predicted annual personal exposure from outdoor sources with surrogate exposure estimates						
Nearest monitor (annual)	0.11	0.05	–0.07	0.02	0.17	0.17
STEAM (2009–13) for NO_2, $PM_{2.5}$, and O_3/ELAPSE (2010) for BC	0.21	0.09	–0.11	–0.15	0.00	–0.14
Temporally aligned STEAM/ELAPSE (2015–19)	0.21	0.00	–0.07	0.05	–0.03	–0.19
LHEM-adjusted STEAM (2009–13)	0.20	0.04	–0.02	Not applicable	Not applicable	Not applicable
Spearman correlation coefficient for personal exposure from outdoor sources during the measurement period with surrogate exposure estimates						
Nearest monitor (annual)	0.10	0.05	0.40	–0.07	0.09	–0.28
STEAM (2009–13) for NO_2, $PM_{2.5}$, and O_3/ELAPSE (2010) for BC	0.23	0.01	0.00	–0.13	–0.01	–0.20
Temporally aligned STEAM/ELAPSE (2015–19) [1]	0.25	0.02	–0.16	0.09	0.07	0.04
LHEM-adjusted STEAM (2009–13)	0.20	–0.06	–0.24	Not applicable	Not applicable	Not applicable

Measurement period and number of participants (n): COPE: June 2015 to October 2017, $n = 75$. BLW: March to May 2019, $n = 162$. PASTA: April 2015 to March 2016, $n = 38$. DEMiST: February 2018 to January 2019, $n = 66$.

(Supplementary Figs B1, B6, B11, and B16). The correlation coefficients between the personal and surrogate exposures for NO_2 were higher, i.e. ≤0.35, when restricted to the period during which participants were indoors (Supplementary Table B2). The correlation coefficients between surrogates varied from –0.17 for the STEAM-modelled estimates and LHEM-adjusted nearest-monitor measurements to >0.9 for the modelled or measured concentrations and their corresponding LHEM adjustments (Supplementary Table B3). For $PM_{2.5}$ in COPE and BLW, the estimated correlations were low, particularly with the STEAM model (ranging from –0.06 to 0.09), except for BLW and the nearest monitor during the measurement period ($r = 0.40$) (Table 2). This may reflect the smaller spatial variability of $PM_{2.5}$ compared with NO_2. For O_3, personal exposures from outdoor sources were not correlated with the nearest-monitor measurements, for neither the study period nor the annual extrapolation, and were inversely correlated with the STEAM-model estimates. However, the correlation became slightly positive when STEAM estimates were temporally aligned ($r = 0.05$). Correlations between surrogates were also relatively low (Supplementary Table B9). The annual estimates were not well correlated with the measurement-period exposures, which may be attributed to the high-O_3 seasonality. Similar patterns were observed for BC, with low Spearman correlations between personal exposure and all surrogate metrics, except the nearest-monitor values for both PASTA and DEMiST ($r = 0.17$). This shows that the nearest-monitor value may better reflect people's movement around their area of residence or work for the DEMiST professional drivers. An increase in the correlations between the personal and ELAPSE model estimates was observed when people were indoors (0.47/0.42 for PASTA/DEMiST, Supplementary Table B11). Graphical assessment of the exposure associations showed no substantial deviation from linearity. A full set of correlations between every exposure estimate, as well as scatter plots per pollutant and study, is provided in Supplementary Section B. When the impact of the exposure separation in the analysis is explored, monthly infiltration efficiencies show low variability across the year (Supplementary Fig. A1) and the Spearman correlations between the surrogate measures and total personal exposure are very similar to those for personal exposure from outdoor sources (Supplementary Table A4).

The mean differences between the predicted personal and surrogate exposures were large when the surrogate approximated the residential outdoor concentrations (i.e. nearest monitor or STEAM/ELAPSE). The difference is substantially lower for LHEM-adjusted estimates, which account for time–activity and infiltration, and indirectly estimate personal exposures.

We also calculated descriptive statistics of the exposure errors by using data from COPE (for NO_2, $PM_{2.5}$, and O_3) and PASTA (for BC) (Table 3). For all pollutants and surrogates except for ELAPSE temporally aligned BC, the errors were positive, i.e. surrogates were higher than personal exposure. For NO_2, the variances of the surrogate estimates were higher compared with the variance of the predicted annual personal exposure from outdoor sources, and the variance ratios ranged from 8.9 for STEAM LHEM-adjusted to 165.0 for STEAM. The Pearson correlation coefficients ranged from 0.08 to 0.25, and the error types were predominantly of classical type (>96% in the mixture). The estimated errors for $PM_{2.5}$ were mixed and the Berkson component was generally larger, from 43% for the nearest-monitor value to 81% for the LHEM-adjusted STEAM estimates, with corresponding variance ratios of 1.3 and 0.2, respectively. The errors for O_3 were high in magnitude, from 29.7 to 61.2 μg/m^3, with a low Pearson correlation between surrogates and personal exposures, high variance ratios (from 35.3 to 52.6), and large classical proportions in the mixture (>94%). The BC errors were predominantly of Berkson type for ELAPSE (temporally aligned or not) with low correlation coefficients and variance ratios, while, for the nearest monitor, the variance ratio was 2.9, which resulted in a 73.8% classical error. Finally, the Pearson correlation coefficients between the $PM_{2.5}$ and NO_2 errors, the $PM_{2.5}$ and O_3 errors, and the NO_2 and O_3 errors ranged from 0.01 to 0.10, from –0.13 to –0.01, and from –0.80 to –0.04, respectively, in COPE (Supplementary Table B13).

Determinants of exposure ME

In our analysis, 54% of the COPE and 58% of the PASTA participants were female and the mean (SD) age was 70.7 (7.9) and 33.8 (9.7), respectively (Table 4). The PASTA participants (younger healthy adults) spent more time out of the home [6.9 (3.0) hours/day on average] compared with the COPD patients [1.5 (1.3) hours/day] and were physically more active. The two groups had similar socio-economic status and the COPD severity of the COPE participants was mild or moderate for 59% of the participants. Most of the healthy adults (in PASTA) were worried about air pollution

Table 3. Magnitude of exposure MEs defined as the difference (in $\mu g/m^3$) between predicted annual personal exposure from outdoor sources and various surrogate exposure assessment methods, correlation coefficients and variance ratios between surrogate measures and personal exposure, and percentages of classical and Berkson error, for the COPE (for NO_2, $PM_{2.5}$, and O_3) and PASTA (for BC) studies.

Pollutant	Surrogate exposure	Nugget	Variance (personal exposure)	Mean difference (surrogate – personal in $\mu g/m^3$)	Pearson correlation (surrogate, personal)	Variance ratio (surrogate/ personal)	% Classical error	% Berkson error
NO_2	STEAM	0.020	2.118	35.61	0.156	165.02	100.0	0.0
	STEAM, LHEM-adjusted			8.99	0.251	8.88	97.0	3.0
	STEAM temporally aligned			28.66	0.126	103.11	100.0	0.0
	Nearest monitor annual			40.00	0.129	154.29	100.0	0.0
	Nearest monitor LHEM-adjusted			11.19	0.079	20.04	96.8	3.2
$PM_{2.5}$	STEAM	0.096	3.922	10.07	0.081	0.53	33.4	66.6
	STEAM, LHEM-adjusted			3.84	0.036	0.25	18.8	81.2
	STEAM temporally aligned			5.74	–0.011	0.29	22.5	77.5
	Nearest monitor annual			7.06	0.104	1.30	57.3	42.7
	Nearest monitor LHEM-adjusted			2.04	0.092	0.49	31.5	68.5
O_3	STEAM	0.008	1.201	57.07	–0.237	42.77	94.6	5.4
	STEAM temporally aligned			61.66	–0.062	35.29	96.3	3.7
	Nearest monitor annual			29.66	–0.002	52.55	98.1	1.9
BC	ELAPSE	0.014	0.971	0.75	0.053	0.07	5.3	94.7
	ELAPSE temporally aligned			–0.44	0.069	0.03	1.6	98.4
	Nearest monitor annual			2.55	–0.021	2.89	73.8	26.2

Table 4. Descriptive statistics for potential exposure ME determinants in two exposure measurement campaigns (COPE and PASTA).

	COPE ($n = 71$)	PASTA ($n = 38$)
Sex [n (%)]		
Male	33 (46)	16 (42)
Female	38 (54)	22 (58)
Age (mean years ±SD)	70.7 ± 7.9	33.8 ± 9.7
Daily time spent out of home (mean hours±SD)	1.5 ± 1.3	6.9 ± 3.0
IMD score (mean ±SD)	22.4 ± 13.2	22.3 ± 10.5
Residence type [n (%)]		
Flat/apartment	34 (48)	12 (32)
Detached/semi-detached	21 (30)	13 (34)
Other[a]	16 (22)	13 (34)
Cooking fuel [n (%)]		
Electric	27 (38)	Not available
Gas	31 (44)	Not available
Other[b]	13 (18)	Not available
COPD severity [n (%)]		
Mild	10 (14)	Not applicable
Moderate	32 (45)	Not applicable
Severe	19 (27)	Not applicable
Very severe	10 (14)	Not applicable
Smoking status [n (%)]		
Never smoked	Not available	31 (82)
Former smoker	Not available	7 (18)
Current smoker	Not available	0 (0)
BMI (mean ± SD)	Not available	22.5 ± 3.0
Time spent physically active (% of study period; mean ± SD)	Not available	29.1 ± 7.6
Self-reported worry of air pollution [n (%)]		
Not worried	Not available	14 (37)
Worried	Not available	24 (63)

[a] Other included bungalows, cottages, terraced houses, and mobile homes.
[b] Other included mix of gas hob and electric oven and wood burning.

in their area (63%). We included 70, 66, and 61 COPE participants in this analysis for NO_2, $PM_{2.5}$, and O_3, respectively, out of 71 due to missing values. No PASTA participants were excluded.

Fig. 2 shows the regression coefficients for all potential determinants of ME (after error standardization) for COPE, NO_2 (Fig. 2a), $PM_{2.5}$ (Fig. 2b), O_3 (Fig. 2c), and PASTA BC (Fig. 2d). Generally, across all pollutants, the type of housing, time spent outdoors, deprivation, and COPD severity (for COPE) were associated with the magnitude of ME, but not all coefficients reached the nominal level of statistical significance ($P < .05$). A detailed description of the error-determinant findings is presented in Supplementary Section C.

Discussion

In this study, we presented multiple surrogate exposure assessment methods that were estimated for the participants of personal exposure measurement campaigns in London. We compared those with the predicted annual personal exposures from outdoor sources estimated based on personal exposure measurements, which was assumed as ‘gold-standard’ exposure. Both gaseous (NO_2, O_3) and particle ($PM_{2.5}$, BC) pollution data were collected, and a common methodology was applied, but with pollutant-specific inputs. These methods are applicable to other settings and air pollutants. The surrogate metrics assessed are those often used in epidemiological studies, i.e. measurements from fixed-site monitors and modelled estimates. We also incorporated methods that took time–activity into account, based on a representative survey conducted in London [35]. The ME was larger for ambient measurements or modelled concentrations compared with time–activity-adjusted surrogates. The correlations between the ‘gold-standard’ and surrogate exposures were generally low. This was not due to the separation into outdoor- and indoor-generated exposure or the annual extrapolation methodology. Classical error was more prominent in the concentration estimates of gases, whereas Berkson error better characterized those of particles.

The modelled surrogates were previously validated and performed well, with R^2 values of 0.54 for BC (ELAPSE—with 20 out of 436 Western European sites used for this model located in London and Oxford) [33] and >0.75 for NO_2, $PM_{2.5}$, and O_3 (STEAM) [28]. However, the

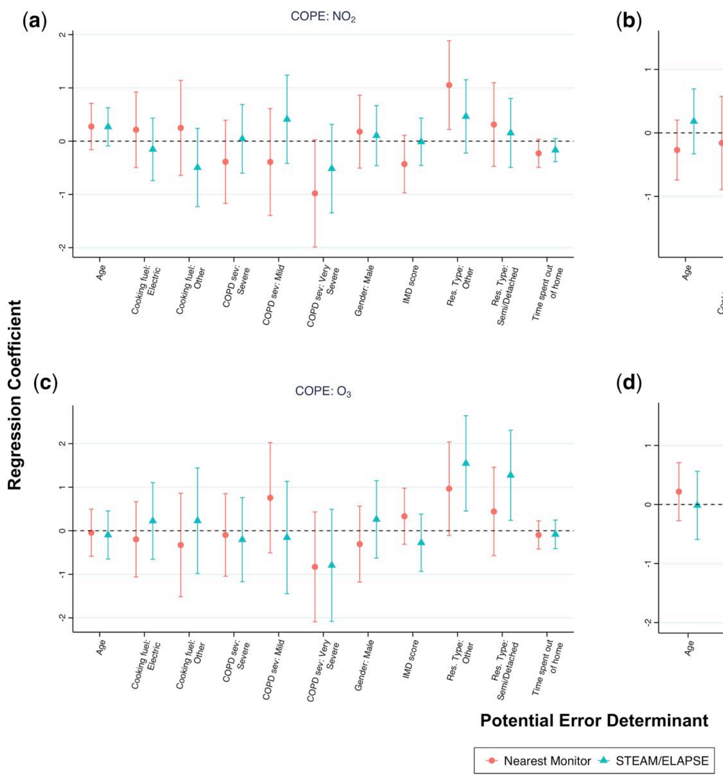

Figure 2. Regression coefficients for potential determinants of ME between standardized predicted annual personal exposure to outdoor sources for COPE participants to NO_2 ($n = 70$), $PM_{2.5}$ ($n = 65$), and O_3 ($n = 60$), as well as PASTA participants ($n = 38$) for BC and the nearest-monitor or modelled (both standardized) surrogates of exposure.

correlations between the surrogate and personal measures were low, exceeding 0.20 only for a few measurement campaigns/pollutant pairs. Part of the differentiation may be attributed to the different characteristics of the participants in the campaigns. For example, we observed a relatively good correlation between personal exposures when participants were indoors and the modelled NO_2 surrogate for COPE (a campaign in which participants tended to spend most of their time indoors). Another part may be that our defined 'gold-standard' exposure, i.e. personal exposure from outdoor sources, also involved several uncertainties in its estimation, increasing the error. The surrogate measures that indirectly adjusted for time–activity information (LHEM) did not lead to increased correlations with personal levels, but scaled down the surrogates and brought them closer to personal exposures. Surrogate exposures that estimated ambient concentrations, e.g. at the residence of a participant, were much higher than the predicted annual or estimated-during-the-measurement-period personal exposure or the LHEM-adjusted estimates. In addition, for BC in PASTA (healthy adults) and DEMiST (professional drivers), we estimated higher correlation coefficients between personal exposures from outdoor sources and the nearest monitor compared with the ELAPSE-modelled estimates, which might reflect the fact that the nearest monitor may better capture some of people's mobility.

We calculated the ME as the difference between the surrogate and personal exposures, and we observed large errors. However, we do not know the extent to which these large errors originated from the uncertainties in calculating surrogate measures or from uncertainties related to the calculation of personal exposure, which was based on the separation of total personal exposure measurements into exposures from outdoor or indoor sources [23]. While it is more evident that static exposure assessment methods, such as monitor measurements or modelled concentrations at the residence, have clear limitations in capturing people's true exposure to pollution, a 'gold-standard' assessment method for individualized exposure to ambient air pollution has not yet been established [4]. In addition, we estimated annual personal exposures, which are more applicable to cohort studies that aim to quantify the impacts of long-term exposure to air pollution. This may have added more uncertainty to our analysis. However, when we tested the sensitivity of our results by restricting the analysis to the measurement period of the campaigns, the correlations between the exposures, as well as the magnitude/variability of the errors, were generally similar.

Often in the literature, smaller MEs than those observed in our study have been reported [3]. Although it is true that some exposure models are more detailed and based on better availability of data, smaller errors are reported when the hypothesized 'gold-standard' exposure is measured under ambient concentrations and the surrogate is modelled at ambient levels at the location of the monitors. Monitors are typically not placed in every location of interest, which can result in preferential sampling, i.e. 'the process that determines the data locations and the process being modelled are stochastically dependent' [38]. The validity of an air-pollution model may be overestimated due to this source of error. Methods that incorporate the differential sampling process can adjust for this [39, 40]. Moreover, if we consider personal exposure from outdoor sources as the 'gold standard', then further uncertainties related to its estimation and different error sources, e.g. time–activity and infiltration, should be considered. Personal exposure from outdoor sources has been investigated in only a few studies [4, 7], but, from an epidemiological and regulatory perspective, it is an important metric that is directly comparable to surrogate exposures that assess ambient pollution levels and both exposures are affected by the same policies [3]. The lower levels of personal exposure observed in this study do not imply that the health impacts of air pollution are smaller than those previously studied. On the contrary, our analysis shows that the large error variability potentially results in large biases in epidemiological analysis, usually underestimating the true impacts.

The low correlation between the personal and surrogate exposures for NO_2 that we observed agree with the results of a previous meta-analysis, which found a pooled correlation of 0.42 for total personal exposure and ambient levels [6]. The season of the study and pre-existing disease of the participants were reported as drivers of these correlations. For $PM_{2.5}$, previous studies showed that both total personal [5] and personal exposure of ambient origin [7] are well correlated with ambient levels, with correlations coefficients of >0.5. Schembari *et al.* (2013) also reported moderate Spearman correlation coefficients between total personal exposures and outdoor levels for $PM_{2.5}$ ($r = 0.4$), PM absorbance ($r = 0.3$), and NO_2 ($r = 0.6$) in a panel study of pregnant women in Barcelona, Spain [41]. Interestingly, the personal exposures in that study were higher than both the outdoor and the indoor levels. No previous studies assessing the correlation between personal O_3 exposure and ambient levels were identified.

Here, we assumed that the errors observed were of additive, mixed (classical and Berkson) type, by using the latent variable mixed error model [36]. Gaseous pollutants (NO_2 and O_3) were found to be prone to classical error, i.e. >94% in the mixture, whereas, for $PM_{2.5}$ and BC, the Berkson component was dominant. Berkson error is not expected to result in bias in epidemiological analysis and our findings suggest that NO_2 and O_3 might be more susceptible to ME bias, which tends to be downwards [8, 9]. This could probably explain the null findings in epidemiological studies of O_3 reported in the literature or contradicting findings from different epidemiological studies, some of which support harmful and others protective effects of NO_2 or O_3.

ME corrections should be applied in air-pollution epidemiological studies, considering the surrogate exposure used and the potential amount of classical error associated with it. Correction methods rely on the existence of a 'gold-standard' exposure assessment, which might not always be available. Previous reviews have shown that <10% of published air-pollution epidemiological studies used methods to address exposure ME, probably due to the lack of personal exposure validation data [14, 15]. However, the studies that have performed corrections showed that biases in the associations between health outcomes and both short- and long-term exposures can substantially underestimate the true exposure–response functions [8, 12, 42, 43].

We investigated whether ME is associated with specific participant characteristics and found that time spent outdoors, type of housing, and deprivation are potential determinants. Living in specific house types, such as terraced houses and/or social houses (also related to deprivation), may be associated with certain air-pollution sources, e.g. gas-boiler exhausts outside windows. Time spent outdoors was associated with MEs of traffic-related pollutants (NO_2/BC). Deprivation was associated with $PM_{2.5}$ error. Those of lower socio-economic status may have larger exposure errors, which leads to larger underestimation of their health impacts. Identifying the characteristics associated with exposure ME is important in epidemiological analysis for better exposure assessment, but also for confounding control. These issues have important implications for policy and future research ME determinants is warranted. We quantified the exposure errors after standardizing the personal and surrogate exposures to remove the systematic difference due to the difference in scale. This approach equalized the exposure variances, which has an impact in the estimation of the classical and Berkson error components in the error mixture and, in turn, in health-effect estimation. Future epidemiological analyses that aim to correct for ME should not perform exposure standardization prior to correction. The difference in scale is taken into account when applying ME correction methods, such as regression calibration [36, 42].

Our study has some strengths. It is based on some of the longest exposure measurement campaigns (>10 000 person-days) and includes samples of individuals with varying characteristics. We assessed exposures to pollutants from ambient sources as 'gold-standard' exposure, which is more reasonable when the surrogates are estimates of outdoor concentrations. Among the surrogate measures, the monitoring network in London is unique in density worldwide, with 42, 14, and 12 monitors for NO_2, $PM_{2.5}$, and O_3 used for this analysis (but only 2 for BC). Additionally, we used a model that adjusted for typical time–activity patterns and the concentrations of indoor microenvironments.

However, there are also limitations. The 'true' exposure of interest (personal exposure from outdoor sources) cannot be directly measured and its estimation, even when based on a large database, entails assumptions and modelling (annual extrapolation), which add uncertainty. We linked our personal exposures with multiple surrogates. We acknowledge that there might be exposure assessment models with higher predictive accuracy than our models, but both the STEAM- and ELAPSE-modelled estimates had relatively high R^2 values. This high accuracy cannot be directly translated into predictions of personal exposures, as the inputs of a personal exposure model would be different from those of ambient air-pollution models, and would include, among other things, the time–activity patterns of the individuals and other personal characteristics, as well as the infiltration efficiencies of buildings [4]. Indirect adjustments for personal exposure that combine time–activity data with ambient pollution modelling have been shown to have little impact in epidemiological analyses [44], but a direct assessment from a personal exposure model would provide further insights. Moreover, the spatial resolution of the models was fine, i.e. postcode-level (with a median of only 14 households per postcode) and $100\,m^2$, respectively. Our measurement campaigns did not include residential outdoor measurements with fixed monitors, which enhanced the uncertainty related to residential outdoor concentrations. Finally, the measurement campaigns included persons with specific health, age, and occupational characteristics, which rendered them less representative of the general population.

In conclusion, exposure ME in air-pollution epidemiology is an important and neglected topic, leading to biased effect estimates. ME bias has been investigated in previous studies, but under a limited range of scenarios for the error magnitude and variability [8–10]. Our study shows the difficulty in assessing ME, given the different exposure assessment methods used in epidemiological analysis and its potentially large magnitude. However, by identifying these difficulties, our study can improve the design of future research. Cohort studies should aim to perform more detailed exposure assessment methods through personal monitoring campaigns. The collection of long-term personal measurements for large populations can be challenging, but carefully designed validation subsamples can provide insight for exposure ME corrections. We further found some evidence of personal behaviours, socio-economic status, and household characteristics that were associated with ME. This shows that there might be variables that are not usually included in epidemiological analysis that need to be collected or that ME bias may affect individuals with specific characteristics. Inclusion of these variables in epidemiological models may provide an indirect adjustment for exposure error when validation data cannot be used. Further studies should be designed to explore the error types, size, and determinants, to allow better-informed epidemiological analysis and correction methods for the estimation of air-pollution effects on health.

Ethics approval

This analysis was performed by using data that had already been collected for previous studies that had obtained ethics approval, i.e. COPE [Camden and Islington Research Ethics Committee (ref. 14/LO/2216)], PASTA [Imperial College Research Ethics Committee (London) on 20 November 2014], DEMiST [KCL BDM Research Ethics Subcommittee (ref HR-16/17-4415)], and BLW [King's College London BDM Research Ethics Subcommittee (ref. RESCM-18/19-9017)].

Acknowledgements

The research described in this article was conducted under contract to the Health Effects Institute (HEI)—an organization jointly funded by the United States Environmental Protection Agency (EPA) (Assistance Award CR 83998101) and certain motor vehicle and engine manufacturers. The contents of this article do not necessarily reflect the views of the HEI or its sponsors, nor do they necessarily reflect the views and policies of the EPA or motor vehicle and engine manufacturers. The HEI retains a non-exclusive licence to publish material from work funded by the HEI. D.E., D.W., B.B., and H.W. were part-funded by the NIHR HPRU in Environmental Exposures and Health—a partnership between the UK Health Security Agency and Imperial College London. D.E. was also funded by the MRC Centre for Environment and Health, Imperial College London. We would like to thank the participants, PIs, and study teams of each of the studies contributing personal measurement datasets for their hard work in data collection and for sharing the datasets with us.

Author contributions

D.E.: writing—original draft, supervision, methodology, investigation, funding acquisition, formal analysis, data curation, conceptualization. D.W.: writing—review and editing, resources, methodology, investigation, formal analysis, data curation. B.B.: writing—review and editing, supervision, investigation, formal analysis, conceptualization. H.Z.: writing—review and editing, methodology, investigation, formal analysis, data curation. A.d.N.: writing—review and editing, methodology, investigation, funding acquisition. S.B.: writing—review and editing, methodology, investigation, funding acquisition. B.K.B.: writing—review and editing, supervision, methodology, funding acquisition, conceptualization. E.S.: writing—review and editing, supervision, methodology, funding acquisition, conceptualization. J.S.: writing—review and editing, supervision, methodology, funding acquisition, conceptualization. K.d.H.: writing—review and editing, methodology, investigation, data curation. K.D.: writing—review and editing, methodology, investigation, data curation. H.W.: writing—review and editing, methodology, funding acquisition, conceptualization. K.K.: writing—review and editing, supervision, methodology, investigation, funding acquisition, formal analysis, conceptualization.

Supplementary data

Supplementary data is available at *IJE* online.

Conflict of interest

B.K.B. owns shares in Royal Dutch Shell and in Scottish and Southern Energy, and her spouse is in receipt of a Shell pension. The other authors have no conflict of interest to declare.

Funding

Health Effects Institute (HEI) funded under agreement #4974-RFA19-1/20–8-3.

Data availability

The summary statistics presented in graphs along with additional information are available in the Supplementary Material. This research involves personal data that cannot be made publicly available. The code used for the analysis is available upon request.

Use of artificial intelligence (AI) tools

No AI tools were used for the design, analysis, and writing of this manuscript.

References

1. Andersen ZJ, Gehring U, De Matteis S *et al.* Clean air for healthy lungs–an urgent call to action: European Respiratory Society position on the launch of the WHO 2021 Air Quality Guidelines. *Eur Respir J* 2021;**58**:2102447.
2. HEI Panel on the Health Effects of Long-Term Exposure to Traffic-Related Air Pollution. Systematic Review and Meta-analysis of Selected Health Effects of Long-Term Exposure to Traffic-Related Air Pollution. 2022. Special Report 23.
3. Katsouyanni K, Evangelopoulos D. Invited perspective: impact of exposure measurement error on effect estimates—an important and neglected problem in air pollution epidemiology. *Environ Health Perspect* 2022;**130**:71302.
4. Evangelopoulos D, Katsouyanni K, Keogh RH *et al.* PM2. 5 and NO2 exposure errors using proxy measures, including derived personal exposure from outdoor sources: A systematic review and meta-analysis. *Environ Int* 2020;**137**:105500.

5. Avery CL, Mills KT, Williams R *et al.* Estimating error in using ambient PM2. 5 concentrations as proxies for personal exposures. *Epidemiology* 2010;**21**:215–23.
6. Meng Q, Svendsgaard D, Kotchmar D, Pinto J. Associations between personal exposures and ambient concentrations of nitrogen dioxide: a quantitative research synthesis. *Atmos Environ* 2012; **57**:322–9.
7. Kioumourtzoglou M-A, Spiegelman D, Szpiro AA *et al.* Exposure measurement error in PM 2.5 health effects studies: a pooled analysis of eight personal exposure validation studies. *Environ Health* 2014;**13**:2–11.
8. Butland BK, Samoli E, Atkinson RW, Barratt B, Katsouyanni K. Measurement error in a multi-level analysis of air pollution and health: a simulation study. *Environ Health* 2019;**18**:1–10.
9. Zeger SL, Thomas D, Dominici F *et al.* Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environ Health Perspect* 2000;**108**:419–26.
10. Wei Y, Qiu X, Yazdi MD *et al.* The impact of exposure measurement error on the estimated concentration–response relationship between long-term exposure to PM 2.5 and mortality. *Environ Health Perspect* 2022;**130**:77006.
11. Chang HH, Fuentes M, Frey HC. Time series analysis of personal exposure to ambient air pollution and mortality using an exposure simulator. *J Expo Sci Environ Epidemiol* 2012;**22**:483–8.
12. Evangelopoulos D, Katsouyanni K, Schwartz J, Walton H. Quantifying the short-term effects of air pollution on health in the presence of exposure measurement error: a simulation study of multi-pollutant model results. *Environ Health* 2021;**20**:1–13.
13. World Health Organisation. *WHO Global Air Quality Guidelines: Particulate Matter (PM2.5 and PM10), Ozone, Nitrogen Dioxide, Sulphur Dioxide and Carbon Monoxide*. Geneva, Switzerland: World Health Organisation, 2021.
14. Samoli E, Butland BK. Incorporating measurement error from modeled air pollution exposures into epidemiological analyses. *Curr Environ Health Rep* 2017;**4**:472–80.
15. Shaw PA, Deffner V, Keogh RH, M.T.G *et al.* Epidemiologic analyses with error-prone exposures: review of current practice and recommendations. *Ann Epidemiol* 2018;**28**:821–8.
16. Moore E, Chatzidiakou L, Jones RL *et al.* Linking e-health records, patient-reported symptoms and environmental exposure data to characterise and model COPD exacerbations: protocol for the COPE study. *BMJ Open* 2016;**6**:e011330.
17. Gerike R, de Nazelle A, Nieuwenhuijsen M; PASTA Consortium *et al.* Physical Activity through Sustainable Transport Approaches (PASTA): a study protocol for a multicentre project. *BMJ Open* 2016;**6**:e009924.
18. Varaden D, Leidland E, Lim S, Barratt B. "I am an air quality scientist"–Using citizen science to characterise school children's exposure to air pollution. *Environ Res* 2021;Oct 1**201**:111536.
19. Lim S, Barratt B, Holliday L, Griffiths CJ, Mudway IS. Characterising professional drivers' exposure to traffic-related air pollution: evidence for reduction strategies from in-vehicle personal exposure monitoring. *Environ Int* 2021;**153**:106532.
20. Chatzidiakou L, Krause A, Popoola OA *et al.* Characterising low-cost sensors in highly portable platforms to quantify personal exposure in diverse environments. *Atmos Meas Tech* 2019; **12**:4643–57.
21. Dons E, Laeremans M, Orjuela JP *et al.* Transport most likely to cause air pollution peak exposures in everyday life: Evidence from over 2000 days of personal monitoring. *Atmos Environ* 2019; **213**:424–32.
22. Kelly F, Armstrong B, Atkinson R; HEI Health Review Committee *et al.* The London low emission zone baseline study. *Research Report (Health Effects Institute)* 2011;**163**:3–79.
23. Zhang H, Evangelopoulos D, Wood D *et al.* Estimating exposure to pollutants generated from indoor and outdoor sources within vulnerable populations using personal air quality monitors: A London case study. *Environ Int* 2025;**198**:109431.
24. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *jair* 2002;**16**:321–57.
25. Diffey BL. An overview analysis of the time people spend outdoors. *Br J Dermatol* 2011;**164**:848–54.
26. Olds T, Maher C, Dumuid D. Life on holidays: differences in activity composition between school and holiday periods in Australian children. *BMC Public Health* 2019;**19**:450–8.
27. Carslaw DC, Ropkins K. Openair—an R package for air quality data analysis. *Environmental Modelling & Software* 2012; **27**:52–61.
28. Dimakopoulou K, Samoli E, Analitis A *et al.* Development and evaluation of spatio-temporal air pollution exposure models and their combinations in the Greater London Area, UK. *Int J Environ Res Public Health* 2022;**19**:5401.
29. Beevers SD, Kitwiroon N, Williams ML, Carslaw DC. One way coupling of CMAQ and a road source dispersion model for fine scale air pollution predictions. *Atmos Environ (1994)* 2012; **59**:47–58.
30. Butland BK, Samoli E, Atkinson RW *et al.* Comparing the performance of air pollution models for nitrogen dioxide and ozone in the context of a multilevel epidemiological analysis. *Environ Epidemiol* 2020;**4**:e093.
31. Samoli E, Butland BK, Rodopoulou S *et al.* The impact of measurement error in modeled ambient particles exposures on health effect estimates in multilevel analysis: a simulation study. *Environ Epidemiol* 2020;**4**:e094.
32. Office for National Statistics (ONS). *Postcode Resident and Household Estimates*. England and Wales: Census 2021, 2021. https://www.nomisweb.co.uk/sources/census_2021_pc (10 July 2024, date last accessed).
33. De Hoogh K, Chen J, Gulliver J *et al.* Spatial PM2. 5, NO2, O3 and BC models for Western Europe–Evaluation of spatiotemporal stability. *Environ Int* 2018;**120**:81–92.
34. 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.
35. Smith JD, Mitsakou C, Kitwiroon N *et al.* London hybrid exposure model: improving human exposure estimates to NO2 and PM2. 5 in an urban setting. *Environ Sci Technol* 2016; **50**:11760–8.
36. Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. *Measurement Error in Nonlinear Models: A Modern Perspective*. London, United Kingdom: Chapman and Hall/CRC, 2006.
37. Noble S, McLennan D, Noble M *et al.* The English indices of deprivation 2019. CLG Ministry of Housing, Editor London. 2019.
38. Diggle PJ, Menezes R, Su TL. Geostatistical inference under preferential sampling. *Journal of the Royal Statistical Society Series C: Applied Statistics* 2010;**59**:191–232.
39. Kloog I, Coull BA, Zanobetti A, Koutrakis P, Schwartz JD. Acute and chronic effects of particles on hospital admissions in New-England. *PLoS One* 2012;**7**:e34664.
40. Feng Y, Wei Y, Coull BA, Schwartz JD. Measurement error correction for ambient PM2. 5 exposure using stratified regression calibration: effects on all-cause mortality. *Environ Res* 2023; **216**:114792.
41. Schembari A, Triguero-Mas M, de Nazelle A *et al.* Personal, indoor and outdoor air pollution levels among pregnant women. *Atmos Environ* 2013;**64**:287–95.
42. Hart JE, Spiegelman D, Beelen R *et al.* Long-term ambient residential traffic–related exposures and measurement error–adjusted risk of incident lung cancer in the Netherlands cohort study on diet and cancer. *Environ Health Perspect* 2015;**123**:860–6.
43. Bergen S, Sheppard L, Kaufman JD, Szpiro AA. Multipollutant measurement error in air pollution epidemiology studies arising from predicting exposures with penalized regression splines. *J R Stat Soc Ser C Appl Stat* 2016;**65**:731–53.
44. Hoek G, Vienneau D, de Hoogh K. Does residential address-based exposure assessment for outdoor air pollution lead to bias in epidemiological studies? *Environ Health* 2024;**23**:75.

International Journal of Epidemiology, 2026, 55, 1–11
https://doi.org/10.1093/ije/dyaf214
Original Article