

RESEARCH

Open Access



# 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

Dimitris Evangelopoulos<sup>1,2\*</sup> , Klea Katsouyanni<sup>1,3</sup>, Joel Schwartz<sup>4</sup> and Heather Walton<sup>1,2</sup>

## Abstract

**Background:** Most epidemiological studies estimate associations without considering exposure measurement error. While some studies have estimated the impact of error in single-exposure models we aimed to quantify the effect of measurement error in multi-exposure models, specifically in time-series analysis of PM<sub>2.5</sub>, NO<sub>2</sub>, and mortality using simulations, under various plausible scenarios for exposure errors. Measurement error in multi-exposure models can lead to effect transfer where the effect estimate is overestimated for the pollutant estimated with more error to the one estimated with less error. This complicates interpretation of the independent effects of different pollutants and thus the relative importance of reducing their concentrations in air pollution policy.

**Methods:** Measurement error was defined as the difference between ambient concentrations and personal exposure from outdoor sources. Simulation inputs for error magnitude and variability were informed by the literature. Error-free exposures with their consequent health outcome and error-prone exposures of various error types (classical/Berkson) were generated. Bias was quantified as the relative difference in effect estimates of the error-free and error-prone exposures.

**Results:** Mortality effect estimates were generally underestimated with greater bias observed when low ratios of the true exposure variance over the error variance were assumed (27.4% underestimation for NO<sub>2</sub>). Higher ratios resulted in smaller, but still substantial bias (up to 19% for both pollutants). Effect transfer was observed indicating that less precise measurements for one pollutant (NO<sub>2</sub>) yield more bias, while the co-pollutant (PM<sub>2.5</sub>) associations were found closer to the true. Interestingly, the sum of single-pollutant model effect estimates was found closer to the summed true associations than those from multi-pollutant models, due to cancelling out of confounding and measurement error bias.

**Conclusions:** Our simulation study indicated an underestimation of true independent health effects of multiple exposures due to measurement error. Using error parameter information in future epidemiological studies should provide more accurate concentration-response functions.

**Keywords:** Air pollution, Measurement error, Mixture error, Effect transfer, Simulations, PM<sub>2.5</sub>, NO<sub>2</sub>

\* Correspondence: [d.evangelopoulos@imperial.ac.uk](mailto:d.evangelopoulos@imperial.ac.uk)

<sup>1</sup>Environmental Research Group, School of Public Health, Imperial College London, Michael Uren Biomedical Engineering Hub, White City Campus, Wood Lane, W12 0BZ, London, UK

<sup>2</sup>NIHR HPRU in Environmental Exposures and Health, Imperial College London, London, UK

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Introduction

Air pollution is the major environmental factor affecting human health [1, 2]. Many publications report associations of morbidity and mortality outcomes with exposure to air pollutants, such as  $PM_{2.5}$ ,  $NO_2$  and  $O_3$  [3, 4]. It is important to assess the independent effects of each pollutant adjusting for potential confounding by other pollutants in the mixture, as policies to reduce emissions differ by targeted pollutant. Multi-pollutant models are the most commonly used method, and, if correctly specified, provide estimates of the independent effects of each exposure [5–7].

Most studies addressing the associations between pollutants and health, or performing health impact assessments (HIA), do not account for exposure measurement error (ME), although it is a well-known problem [8–11]. ME may lead to biased concentration-response functions (CRFs) for which the magnitude and type of bias is not assessed or corrected. While the methodology to correct for ME is expanding, there is a gap between theory and practice [12, 13]. To evaluate ME we must define the “gold standard” against which “proxy” measures are compared. For exposure to ambient concentrations, the most appropriate gold standard is individual exposure to ambient sources of the pollutant of interest. However, most epidemiological studies investigating ME in air pollution exposure, measure total personal exposure, to which indoor and personally-generated sources contribute substantially [14].

An appropriate framework describing the types of error for exposure variables has not been discussed thoroughly in the literature. Zeger et al. (2000) mention that “classical and Berkson models represent two extremes of a continuum”, so the proper model combines elements of each type [9]. Few studies have examined the impact of mixture error [15–17]. Previous studies examining the impact of ME in regression models either only used single-pollutant models or addressed only spatial heterogeneity of the pollutants and their errors or have made specific assumptions for their error definition using empirical data from certain locations which might not be generalisable [18–23]. Multi-pollutant models, in addition, provide unstable estimates because of the complex relationships between pollutants, both in space and time, and the differing degree and structure of their errors - especially the variability and correlation between errors.

In our study, we focus on multi-pollutant models using a mixture error model of classical and Berkson components based on the error decomposition of Zeger and colleagues (2000). The percentages of each type depend on the pollutant under investigation, i.e. fine particulate matter,  $PM_{2.5}$ , or nitrogen dioxide,  $NO_2$ . We used simulated data to generate a wide range of plausible scenarios of the pollutant errors under a mixture error

model, and to quantify the real impact of ME on the multi-pollutant model effect estimates, specifically in time-series analysis.

## Methods

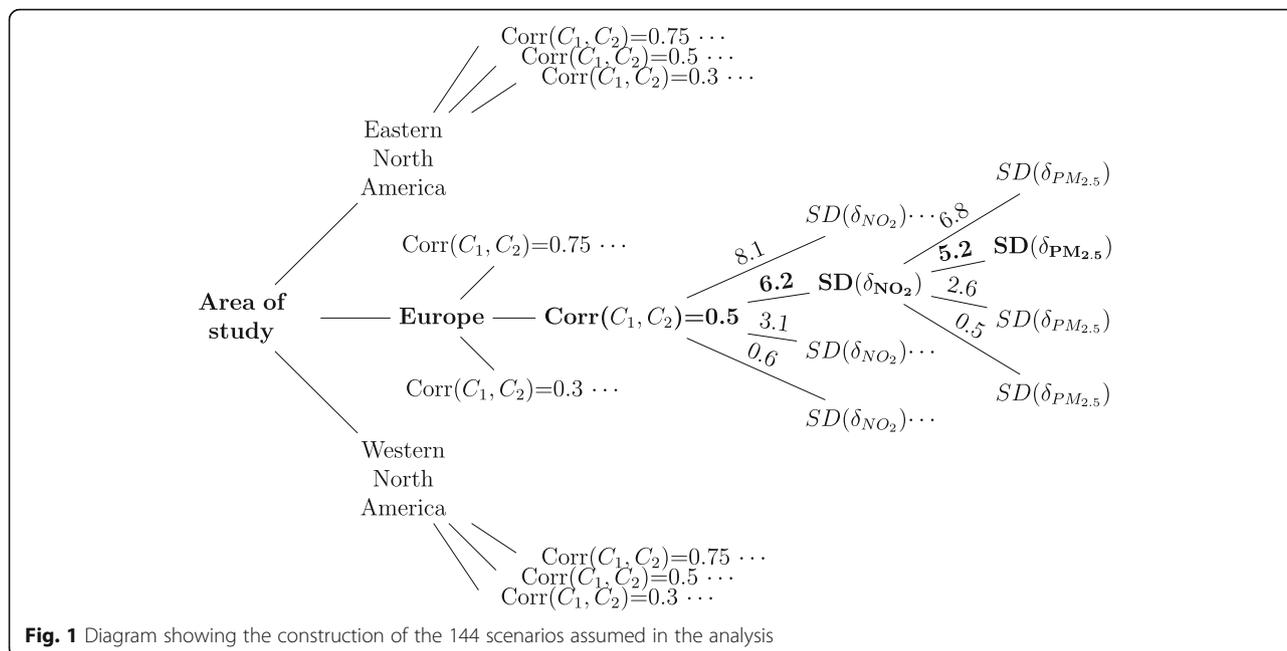
### Simulation set-up

We previously conducted a systematic review on the differences between ambient concentrations and personal exposures from outdoor origins to inform our simulation inputs and quantify the magnitude and variability of the pollutant errors [24]. In particular, we created a simulated daily time-series for error-free (A) and error-prone (C) exposures and a health outcome over a period of 4 years, under various scenarios for factors that drive ME bias (Fig. S1, Supplementary material). Our focus was only on the quantification of exposure measurement error bias, so we did not include some features of time-series that may introduce other forms of bias and need to be taken into account in real data analysis, such as autocorrelation, trends or confounders measured with error.

Each scenario was created as a combination of key parameters driving ME bias, i.e. the correlation between the exposures, the error variability ( $SD(\delta_{PM_{2.5}})$ ,  $SD(\delta_{NO_2})$ ) and the true exposure variance/error variance ratio ( $SD(A_{PM_{2.5}})/SD(\delta_{PM_{2.5}})$ ,  $SD(A_{NO_2})/SD(\delta_{NO_2})$ ) [23]. The latter was informed according to data from three different areas, for which we had adequate information [24]. In total, 144 scenarios were investigated, and their inputs are summarised in Figure 1. Due to space constraints, we expanded only one branch in the tree plot (in bold); this was assumed to be the core scenario compared with which all sensitivity analyses were performed. For each scenario, 1000 simulated datasets were generated.

Briefly, we assumed:

- Variance ratios: We created scenarios separately for Europe, Eastern and Western North America which correspond to different ratios of the true exposure variance over the error variance. Differences in the ratios between these regions can be attributed to various factors, such as the use of Diesel vehicles, air conditioning, and legislation [25].
- Correlation between the exposures: To the best of our knowledge, no data for the correlation between personal exposures from outdoor sources to  $PM_{2.5}$  and  $NO_2$  are reported in the literature. Thus, we assumed that it lies between published correlation coefficients of total personal exposures and correlations of the ambient measurements for the same pollutants [26–28]. A range of values was hypothesized to cover various plausible scenarios.
- Correlations between the errors: Because we did not have real data to estimate the relationship between



**Fig. 1** Diagram showing the construction of the 144 scenarios assumed in the analysis

these errors, an analytical approach for their covariance was applied using the pairwise associations of personal and ambient concentrations for NO<sub>2</sub> and PM<sub>2.5</sub> (eq. 2).

- Error structures: Little is known about the distributions and types of exposure ME. We assumed that the errors are additive on the logarithmic scale [29, 30]. Moreover, we assumed that a mixture of classical and Berkson error best describes the type of exposure misclassification [9]. However, scenarios of entirely classical and entirely Berkson error, alternative figures for the classical-Berkson ratio and multiplicative error were also assessed.

**Input data**

The key input parameters are summarised in Table 1. We assumed different “true” exposures and error variability for each study area which resulted in different exposure variance over error variance ratios. Three correlation coefficients between the exposures (low = 0.25/moderate = 0.5/high = 0.75) and four values for the error variability for each pollutant were hypothesized. For example, for Europe (core scenario) we hypothesized moderate error variability (5.2 µg/m<sup>3</sup> for PM<sub>2.5</sub> and 6.2 ppb for NO<sub>2</sub>), very low (0.5 µg/m<sup>3</sup> and 0.6 ppb respectively (0.1 × moderate)), low (2.6 µg/m<sup>3</sup> and 3.1 ppb respectively (0.5 × moderate)), and high (6.8 µg/m<sup>3</sup> and 8.1 ppb respectively (1.3 × moderate)). The 144 scenarios

tested derived from all the possible combinations of these parameters.

**Generating “error-free” or “true” exposures (A):**

We used a bivariate log-normal distribution to generate 1461 daily measurements of the “error-free” exposures for each scenario **s**, as follows:

$$\text{Lognormal}(\boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s), \boldsymbol{\mu}_s = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \boldsymbol{\Sigma}_s = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{2,1} & \sigma_2^2 \end{bmatrix}$$

where  $\boldsymbol{\mu}_s$  is the vector of pollutant means,  $\boldsymbol{\Sigma}_s$  the variance-covariance matrix and 1, 2 = NO<sub>2</sub> and PM<sub>2.5</sub>, respectively. For example, for the European area, we had:

$$\boldsymbol{\mu}_{eu} = \begin{bmatrix} 21.6 \\ 21.1 \end{bmatrix}, \boldsymbol{\Sigma}_{eu} = \begin{bmatrix} 9.0^2 & 27.6 \\ 27.6 & 10.9^2 \end{bmatrix}.$$

**Generating “error-prone” or “apparent” exposures (C):**

We considered a mixture of classical and Berkson ME additive on the logarithmic scale as the most appropriate model because the sources of error can originate from both classical error (omitting individuals’ mobility, the infiltration efficiency of the buildings, and instrument errors) and Berkson error from unincorporated spatial heterogeneity [9]. Various scenarios for the errors allowed evaluation of how the error type affects the health effect estimates, especially in two-pollutant models where the magnitude and direction of bias cannot easily be

**Table 1** Simulation inputs for the assumed “true” exposures and the error variability of the “error-prone” exposures

Area of Study	“True” PM <sub>2.5</sub> - Mean (SD) (µg/m <sup>3</sup> )	“True” NO <sub>2</sub> - Mean (SD) (ppb)	PM <sub>2.5</sub> Error - SD (µg/m <sup>3</sup> )	NO <sub>2</sub> Error - SD (ppb)
Eastern North America	19.0 (8.6)	20.7 (11.6)	5.7	7.3
Europe (core scenario)	21.1 (10.9)	21.6 (8.9)	5.2	6.2
Western North America	18.7 (8.3)	22.4 (10.9)	5.6	7.3

predicted in typical epidemiologic studies without specifically investigating ME.

**Classical error**

Apparent exposures of classical error were generated simply by adding random error to the assumed true exposures. For this type of error, we have:

$$C_i^t = A_i^t + \delta_i^t \tag{1}$$

where i is NO<sub>2</sub> or PM<sub>2.5</sub>, t is day, t = 1,2, ..., 1461, C and A are as defined above, and the errors, δ, of the pollutants were created from a bivariate normal distribution:

$$\begin{aligned} \delta_i^t \mid A_i^t &\sim N(\mu, \Sigma), \mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma \\ &= \begin{bmatrix} \text{Var}(\delta_{NO_2}^t) & \text{Cov}(\delta_{NO_2}^t, \delta_{PM_{2.5}}^t) \\ \text{Cov}(\delta_{NO_2}^t, \delta_{PM_{2.5}}^t) & \text{Var}(\delta_{PM_{2.5}}^t) \end{bmatrix} \end{aligned}$$

The elements in the diagonal of Σ are the error variances as estimated from [24]. The covariance was estimated analytically:

$$\begin{aligned} \text{Cov}(\delta_{NO_2}^t, \delta_{PM_{2.5}}^t) &= \text{Cov}(C_{NO_2}^t - A_{NO_2}^t, C_{PM_{2.5}}^t - A_{PM_{2.5}}^t) \tag{2} \\ &= \text{Cov}(C_{NO_2}^t, C_{PM_{2.5}}^t) - \text{Cov}(C_{NO_2}^t, A_{PM_{2.5}}^t) \\ &\quad - \text{Cov}(A_{NO_2}^t, C_{PM_{2.5}}^t) + \text{Cov}(A_{NO_2}^t, A_{PM_{2.5}}^t) \end{aligned}$$

There was no information in the literature for the associations between C and A, so we used covariances from studies of total personal exposure instead [26–28].

**Berkson error**

In general, a Berkson error model can be described as follows:

$$A_i^t = C_i^t + \delta_i^t, \tag{3}$$

We used a formula to generate Berkson error proposed by Carroll et al. (2006) [11]. Briefly, if C = A + δ (i.e. classical model) we know that the best linear predictor of A given C is (1-λ)·E(A) + λ·C, where  $\lambda = \frac{\sigma_c^2}{\sigma_c^2 + \sigma_\delta^2}$ . Then, the following equation (4) results in the generation of a Berksonian-type error-prone exposure. For simplicity, we omitted the indicators i, t:

$$C = \mu_A + (A - \mu_A) \cdot \frac{\sigma_c^2}{\sigma_c^2 + \sigma_\delta^2} + \delta^* \tag{4}$$

where  $\delta^* \sim N(0, \frac{\sigma_c^2 \cdot \sigma_\delta^2}{\sigma_c^2 + \sigma_\delta^2})$  and  $\sigma_\delta^2$  is the error variance (Supplementary material).

**Mixture error**

Using the error decomposition from Zeger and colleagues [9], with data from other studies [31, 32], we quantified the error-prone variables of the mixture type. We estimated the error as 43% classical and 57% Berkson for PM<sub>2.5</sub> and 33% classical and 67% Berkson for NO<sub>2</sub> in the main analysis, while in sensitivity analyses we assumed increased percentages of classical error for both pollutants, i.e. either (55,45%) for PM<sub>2.5</sub> and (45, 55%) for NO<sub>2</sub> or (70,30%) for PM<sub>2.5</sub> and (60,40%) for NO<sub>2</sub>. As the hypothesized true exposures are log-normally distributed, we assumed additive error on the log-scale for the ME models based on previous studies, but also applied a multiplicative approach [29, 30].

The mixture model consists of the variables described in eqs. 1 and 2, along with a latent intermediate variable L, between A and C, that allows for mixtures of Berkson and classical error as described elsewhere [15, 33]. Briefly, the model is:

$$\begin{aligned} A &= L + \delta_b \\ C &= L + \delta_c \end{aligned}$$

where δ<sub>b</sub>, δ<sub>c</sub> denote Berkson-type and classical-type error (Supplementary material).

**Generating all-cause mortality**

For each day t, the number of deaths Y<sub>t</sub> was assumed to follow a Poisson distribution with overdispersion φ. The mean μ<sub>t</sub> of the Poisson distribution was created to be dependent only on the simulated “true” concentrations of both pollutants on every particular day as follows:

$$\log(\mu_t) = \beta_0 + \beta_1 \cdot A_{NO_2}^t + \beta_2 \cdot A_{PM_{2.5}}^t$$

where A<sup>t</sup><sub>NO<sub>2</sub></sub>, A<sup>t</sup><sub>PM<sub>2.5</sub></sub> are the daily concentrations for NO<sub>2</sub> and PM<sub>2.5</sub>, respectively.

To simulate this Poisson over-dispersed data, we assumed that daily mortality followed a Negative Binomial distribution, NB (μ<sub>t</sub>, θ), where μ<sub>t</sub> is the mean daily number of deaths, and θ = φ/(μ<sub>t</sub>-1). To get

reasonable estimates for  $\beta_0$ ,  $\phi$  we used real mortality and air pollution data during 2011–2014 in Greater London.

For the selection of  $\beta_1$ ,  $\beta_2$  we reviewed the literature for studies that reported associations between short-term exposure to  $\text{NO}_2$  and  $\text{PM}_{2.5}$  and all-cause mortality and derived their CRFs. We chose plausible CRF values from a recent meta-analysis of time series studies despite the fact that they are not adjusted for ME, i.e. a 0.6 and 0.54% increase in all-cause mortality per 10 units increase in  $\text{NO}_2$  and  $\text{PM}_{2.5}$  respectively ( $\beta_1 = 0.0006$ ,  $\beta_2 = 0.00054$ ) [34]. We also tested whether the results remain unchanged when the true coefficients were assumed to be half or twice the above percentage increases or when only one pollutant had a true health effect ( $\beta_1 = 0$  for the co-pollutant).

**Epidemiological analysis**

In the context of this study, we are interested in the ME bias - quantified as the difference between the health effect estimates of error-free and error-prone exposures. The Poisson time-series model allowing for over-dispersion used in every iteration was:

$$\log(E(Y_t)) = \beta_0 + \beta_1 \cdot C_{\text{NO}_2}^t + \beta_2 \cdot C_{\text{PM}_{2.5}}^t$$

where  $Y_t$  is the death count for day  $t$  and  $C_{\text{NO}_2}^t$ ,  $C_{\text{PM}_{2.5}}^t$  the corresponding error-prone exposure based on every scenario. We also calculated coverage probability as the

percentage of 95% confidence intervals that include the assumed true exposure-response association, and power as the percentage of statistically significant estimates at the 5% level.

**Results**

We confirmed that the generated exposure variables had close to the expected mean values (Table 1) and the expected distributions for the classical, Berkson or mixed error types (Fig. S2, Supplementary Material).

Table 2 contains the average Poisson regression estimates across all scenarios by areas under investigation. Pure Berkson error model resulted in overestimation of the true mortality effect for all areas and pollutants by 1.5 to 8.2%, except for the  $\text{NO}_2$  coefficient for Europe (a decrease of 7.9%), the assumed area scenario with the lower exposure variance/error variance ratio and higher corresponding ratio for the co-exposure. Otherwise, the mortality effects were underestimated with greater bias observed in Europe for  $\text{NO}_2$  (27.4%, mixture model). Interestingly, the bias in North America was similar for both pollutants, around 15–19% for classical error and 9–13% for mixture (area scenarios with similar, and relatively low exposure variance/error variance ratios for both pollutants). In Europe, the bias was significantly increased for  $\text{NO}_2$  compared with  $\text{PM}_{2.5}$ , 10.0 and 25.1% for classical and 1.3 and 27.4% for mixture error respectively. Coverage and power for both pollutants across all

**Table 2** Summary of the true and error-prone regression coefficients, their standard errors (SE)  $\times 10^{-4}$  and relative bias from 144,000 simulated datasets on the impact of three error models (classical, Berkson and mixture) on 2-pollutant Poisson regression by area of study. Results presented for all scenarios ( $N = 48,000$  in each row)

Exposure	CRFs <sup>a</sup> : Area	$\text{PM}_{2.5}$ : $\beta_1 = 5.4^a$					$\text{NO}_2$ : $\beta_2 = 6^a$				
		$\hat{\beta}_1$	(SE <sub>W</sub> )/ (SE <sub>B</sub> ) <sup>b</sup>	Bias (%) <sup>c</sup>	Coverage Probability (%)	Power (%)	$\hat{\beta}_2$	(SE <sub>W</sub> )/ (SE <sub>B</sub> ) <sup>b</sup>	Bias (%) <sup>c</sup>	Coverage Probability (%)	Power (%)
True	Europe	5.37	(1.48)/(3.46)	–	–	–	6.00	(1.79)/(4.20)	–	–	–
	East NA	5.42	(1.88)/(4.41)	–	–	–	5.98	(1.39)/(3.25)	–	–	–
	West NA	5.40	(1.95)/(4.58)	–	–	–	5.99	(1.48)/(3.45)	–	–	–
Classical	Europe	4.86	(1.36)/(3.27)	–10.0	55.3	76.0	4.49	(1.52)/(3.82)	–25.1	51.1	67.6
	East NA	4.58	(1.67)/(4.02)	–15.1	54.6	65.2	4.97	(1.25)/(3.09)	–17.2	53.9	80.4
	West NA	4.52	(1.72)/(4.17)	–16.3	54.3	63.7	4.85	(1.30)/(3.24)	–19.2	52.8	77.2
Berkson	Europe	5.84	(1.70)/(4.02)	+8.2	59.3	75.4	5.53	(2.46)/(5.89)	–7.9	58.8	62.2
	East NA	5.78	(2.73)/(6.43)	+7.1	59.3	62.2	6.17	(1.95)/(4.58)	+2.8	59.7	75.6
	West NA	5.62	(3.29)/(7.88)	+4.1	59.3	60.2	6.09	(2.50)/(5.87)	+1.5	59.2	71.1
Mixture <sup>d</sup>	Europe	5.33	(1.49)/(3.57)	–1.3	57.9	76.1	4.36	(2.00)/(5.10)	–27.4	55.0	60.8
	East NA	4.91	(1.91)/(4.61)	–9.1	57.1	63.2	5.37	(1.48)/(3.58)	–10.6	58.0	76.9
	West NA	4.84	(1.99)/(4.81)	–10.3	57.2	61.5	5.20	(1.58)/(3.85)	–13.3	57.1	72.8

<sup>a</sup>Concentration-response functions for generation of the health outcome

<sup>b</sup> SE<sub>W</sub>:Within-simulations (or model-based) standard error, SE<sub>B</sub>:Between-simulations (or empirical) standard error

<sup>c</sup> Relative bias =  $\frac{(\hat{\beta}_i - \beta_i)}{\beta_i}$

<sup>d</sup> (Classical,Berkson) percentages: (43,57%) for  $\text{PM}_{2.5}$ , (33,67%) for  $\text{NO}_2$

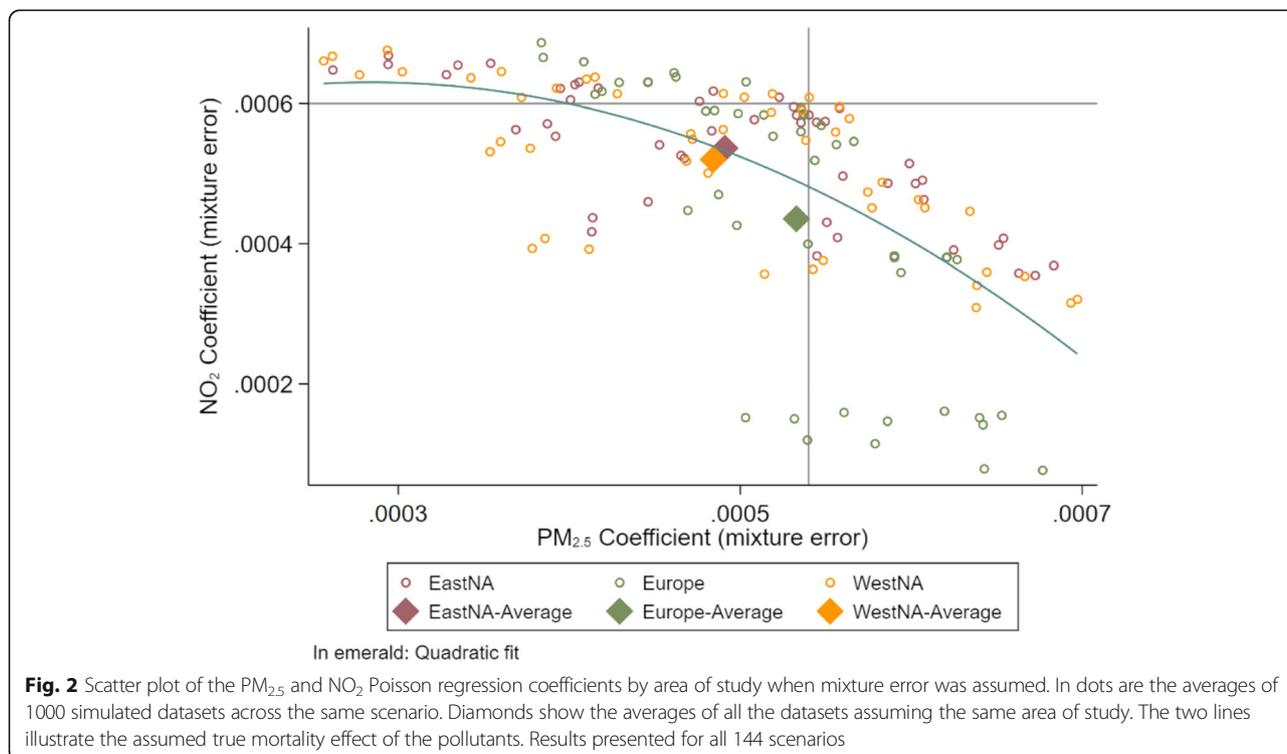
scenarios by area and error type were 51–60% and 60–80% respectively.

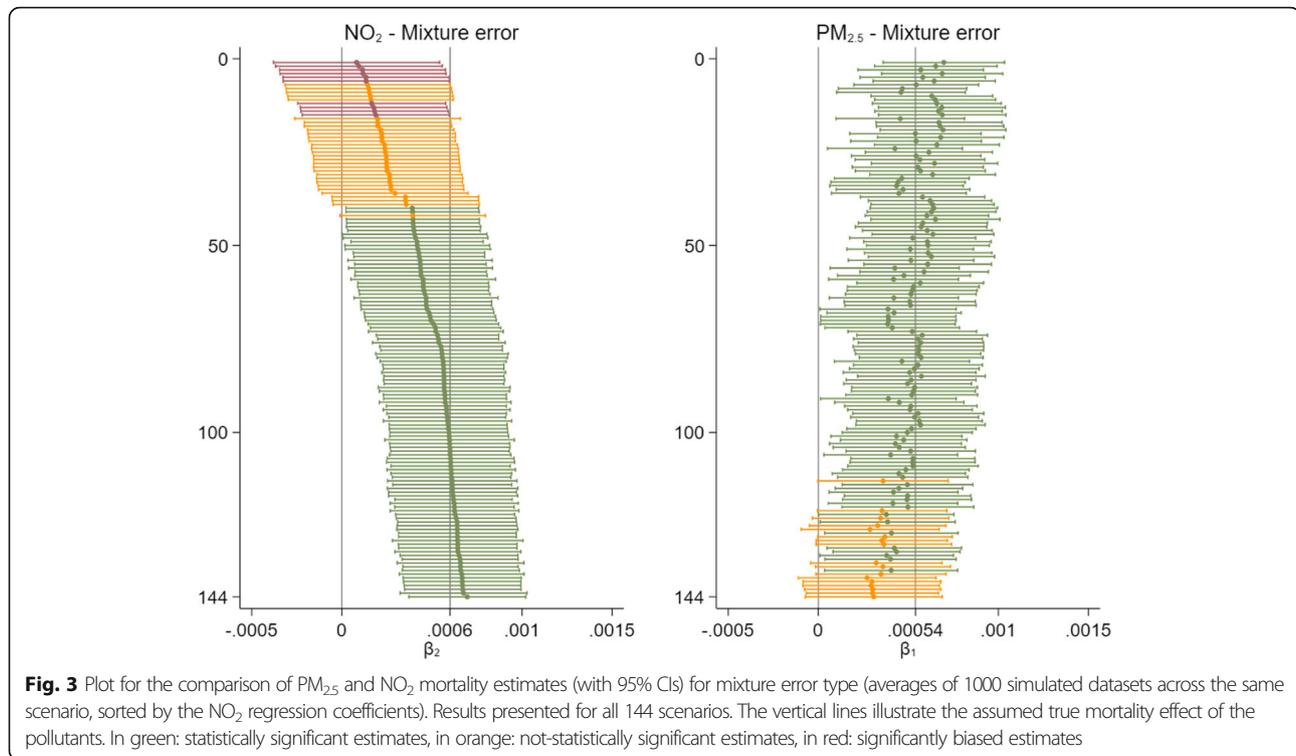
The results of mixture error are illustrated in Fig. 2 showing the relationship between PM<sub>2.5</sub> and NO<sub>2</sub> coefficients by study area for all scenarios. No scenario resulted in overestimation of the effect of both pollutants (upper right quartile). Between pollutants, higher values of  $\hat{\beta}_1$  tend to be associated with lower values of  $\hat{\beta}_2$ , in an approximate quadratic negative relationship. No clear pattern was observed by study area, except for a group of European results in the lower right of the graph (> 70% decrease in the NO<sub>2</sub> effect estimate). NO<sub>2</sub> coefficient attenuation is higher than in other scenarios because the hypothesized exposure variance/error variance ratio was lower compared to PM<sub>2.5</sub>.

Figure 3 plots the mortality estimates for both pollutants sorted in increasing order of NO<sub>2</sub> effect size (left). Effect transfer can clearly be observed either from NO<sub>2</sub> to PM<sub>2.5</sub> (upper part) or from PM<sub>2.5</sub> to NO<sub>2</sub> (lower part), but also, there are some scenarios where both coefficients were attenuated (middle part). Interestingly, in the middle and even at the top of the graph, there are scenarios in which the estimates for both pollutants are to the left of the true beta line i.e. attenuated. Nonetheless, large attenuations for both pollutant estimates did not occur simultaneously, as can be seen from the absence of points in the bottom left corner of Fig. 2.

Table 3 illustrates the model estimates with different assumed values for the correlation between the exposures and the error variances. Results from the most plausible error model are reported, i.e. mixture error. Effect transfer is clearly observed, as the effect estimates for one exposure are decreasing when its error variability is increasing (less precise exposure), while the co-exposure’s coefficients are increasing, and its error variability is kept fixed. For example, in the “moderate” correlation scenarios, keeping the PM<sub>2.5</sub> error variability fixed at “low” ( $SD(\delta_{PM_{2.5}}) = 2.9 \mu\text{g}/\text{m}^3$ ), the bias in the PM<sub>2.5</sub> Poisson regression estimates range from –3.7% in the “very low” NO<sub>2</sub> error variability scenario to +17.8% in the “high” scenario. On the contrary, the corresponding bias for NO<sub>2</sub> varies from +0.7% to –50.0% (“very low” to “high”  $SD(\delta_{NO_2})$ ). Bias in the effect estimates varies highly by assumed error structure from 26.1% overestimation to 44.0% underestimation of the true effect for PM<sub>2.5</sub>, and from +12.1% to –58.4% for NO<sub>2</sub>. The highest biases are observed in the most extreme scenarios, but in the more plausible “low”-“moderate” error variability scenarios, bias is not negligible (values up to approximately 30% towards the null). Correlation between the exposures does not seem to be a driving factor for the bias.

Comparing multi- and single-pollutant model coefficients ( $\beta_M$  and  $\beta_S$  respectively), Table 4 suggests that,





in the absence of ME, using  $\beta_S$  may result in a large overestimation of the combined effect of PM<sub>2.5</sub> and NO<sub>2</sub>. Thus,  $\beta_S$ , in general, includes effects attributable to both pollutants where both pollutants have a true effect on mortality. For mixture error, regression coefficients increased by 22.5% for PM<sub>2.5</sub> and 17.7% for NO<sub>2</sub>, when single- instead of multi-pollutant models were fitted. Surprisingly, the combined effect,  $\beta_{S_1} + \beta_{S_2}$ , of the mixture error-prone variables was closer to the true combined effect, i.e. 5.4 + 6.0, compared to  $\beta_{M_1} + \beta_{M_2}$ . No significant differences were observed between the standard errors of  $\beta_M$  and  $\beta_S$ .

We performed sensitivity analyses with the Europe, additive, mixture error, core scenario to check which of the simulation inputs are driving ME bias. No difference in the relative bias was observed when we assumed a “true” CRF half or twice the core one. Mortality estimates were attenuated even more when only one pollutant had a true effect on health compared to the case where both had an effect. Interestingly, effect transfer is observed from the pollutant with the true effect to the co-pollutant assumed to have no true effect. When the classical part in the error mixture was increased for both pollutants, the relative bias increased for PM<sub>2.5</sub> (from 1.3 to 6.0%), while it only decreased slightly for NO<sub>2</sub> (from 27.4 to 25.4%). Finally, multiplicative error resulted in

great underestimation of the exposure-response relationship, i.e. 84.5 and 90.0% for PM<sub>2.5</sub> and NO<sub>2</sub> respectively (Table 5).

## Discussion

We performed a simulation study to quantify the bias in mortality effect estimates caused by ME in multi-pollutant, time-series models including PM<sub>2.5</sub> and NO<sub>2</sub>. While the impact of ME can be more easily predicted when single exposures are measured with error, multiple error-prone exposures of any error type (i.e. purely classical, purely Berkson or mixture) can distort the health effect estimates. Our results can be applied to other outcomes and exposures as well.

Mixture error model was found to attenuate the effect of both pollutants, with higher attenuation for NO<sub>2</sub>, the exposure variable measured with more error as found previously [24]. ME also reduced coverage of 95% confidence intervals and statistical power. The largest underestimations of the true effect were in North America for PM<sub>2.5</sub> (10.3%) and in European studies for NO<sub>2</sub> (27.4%). Differences by study area were observed because of the different pollutant variability/error variability ratios in these areas and/or potentially other unobserved parameters related to ME. For example, the exposure assessment method used (e.g. measured or modelled

**Table 3** Summary of the regression coefficients, their standard errors (SE)( $\times 10^{-4}$ ) and relative bias of 144,000 simulated datasets on the impact of mixture error model on 2-pollutant Poisson regression by the error variability of PM<sub>2.5</sub> and NO<sub>2</sub>. Results presented for all scenarios ( $N = 3000$  in each row)

CRFs <sup>a</sup> :			PM <sub>2.5</sub> : $\beta_1 = 5.4^a$			NO <sub>2</sub> : $\beta_2 = 6^a$		
Correlation between exposures	PM <sub>2.5</sub> error variability <sup>b</sup>	NO <sub>2</sub> error variability <sup>b</sup>	$\hat{\beta}_1$	(SE <sub>W</sub> )/(SE <sub>B</sub> ) <sup>c</sup>	Bias (%) <sup>d</sup>	$\hat{\beta}_2$	(SE <sub>W</sub> )/(SE <sub>B</sub> ) <sup>c</sup>	Bias (%) <sup>d</sup>
Low	Very low	Very low	5.44	(1.76)/(4.15)	+ 0.7	5.90	(1.54)/(3.62)	-1.7
		Low	5.49	(1.76)/(4.00)	+ 1.6	5.74	(1.58)/(3.66)	-4.3
		Moderate	6.10	(1.74)/(4.12)	+ 13.0	4.45	(1.71)/(4.02)	-25.8
		High	6.57	(1.73)/(4.08)	+ 21.7	2.85	(1.88)/(4.56)	-52.5
	Low	Very low	5.20	(1.77)/(4.23)	-3.7	6.04	(1.54)/(3.62)	+ 0.7
		Low	5.22	(1.76)/(4.16)	-3.4	5.69	(1.57)/(3.72)	-5.1
		Moderate	5.88	(1.75)/(4.11)	+ 8.9	4.56	(1.70)/(4.17)	-23.9
		High	6.36	(1.74)/(4.05)	+ 17.8	3.00	(1.88)/(4.56)	-50.0
	Moderate	Very low	4.31	(1.81)/(4.30)	-20.2	6.35	(1.53)/(3.58)	+ 5.8
		Low	4.58	(1.80)/(4.29)	-15.3	6.10	(1.56)/(3.78)	+ 1.7
		Moderate	4.94	(1.79)/(4.31)	-8.5	4.88	(1.70)/(4.02)	-18.7
		High	5.42	(1.78)/(4.21)	+ 0.3	3.16	(1.88)/(4.59)	-47.4
High	Very low	3.46	(1.85)/(4.31)	-36.0	6.64	(1.52)/(3.57)	+ 10.6	
	Low	3.72	(1.85)/(4.40)	-31.1	6.37	(1.56)/(3.67)	+ 6.1	
	Moderate	4.07	(1.84)/(4.48)	-24.6	5.16	(1.69)/(4.05)	-14.0	
	High	4.32	(1.84)/(4.37)	-20.1	3.21	(1.88)/(4.64)	-46.5	
Moderate	Very low	Very low	5.36	(1.77)/(4.19)	-0.8	5.99	(1.56)/(3.71)	-0.2
		Low	5.52	(1.77)/(4.34)	+ 2.2	5.59	(1.60)/(3.69)	-6.9
		Moderate	6.12	(1.76)/(4.19)	+ 13.4	4.39	(1.72)/(4.07)	-26.9
		High	6.78	(1.74)/(4.15)	+ 25.5	2.82	(1.89)/(4.56)	-53.0
	Low	Very low	5.06	(1.79)/(4.17)	-6.4	6.06	(1.56)/(3.62)	+ 0.9
		Low	5.15	(1.78)/(4.23)	-4.6	5.73	(1.59)/(3.79)	-4.5
		Moderate	5.87	(1.76)/(4.07)	+ 8.6	4.53	(1.71)/(4.14)	-24.6
		High	6.46	(1.74)/(4.16)	+ 19.6	3.07	(1.89)/(4.54)	-48.9
	Moderate	Very low	4.20	(1.83)/(4.32)	-22.2	6.30	(1.55)/(3.74)	+ 5.1
		Low	4.31	(1.81)/(4.37)	-20.2	6.17	(1.57)/(3.75)	+ 2.9
		Moderate	4.90	(1.80)/(4.26)	-9.3	5.03	(1.71)/(4.13)	-16.2
		High	5.64	(1.79)/(4.30)	+ 4.4	3.11	(1.88)/(4.66)	-48.2
High	Very low	3.11	(1.86)/(4.42)	-42.3	6.72	(1.54)/(3.60)	+ 12.1	
	Low	3.57	(1.85)/(4.40)	-33.9	6.39	(1.56)/(3.67)	+ 6.5	
	Moderate	4.17	(1.84)/(4.37)	-22.8	5.26	(1.70)/(4.10)	-12.3	
	High	4.55	(1.83)/(4.39)	-15.8	3.17	(1.88)/(4.64)	-47.2	
High	Very low	Very low	5.43	(1.78)/(4.15)	+ 0.5	5.89	(1.57)/(3.74)	-1.9
		Low	5.51	(1.78)/(4.21)	-2.1	5.57	(1.60)/(3.86)	-7.3
		Moderate	6.21	(1.76)/(4.09)	+ 15.0	4.30	(1.73)/(4.09)	-28.3
		High	6.81	(1.74)/(4.08)	+ 26.1	2.49	(1.90)/(4.73)	-58.4
	Low	Very low	4.90	(1.80)/(4.24)	-9.3	6.16	(1.57)/(3.70)	+ 2.7
		Low	4.99	(1.82)/(4.27)	-7.5	5.46	(1.62)/(3.91)	-9.0
		Moderate	5.77	(1.78)/(4.19)	+ 6.8	4.36	(1.73)/(4.11)	-27.3
		High	6.35	(1.76)/(4.12)	+ 17.6	2.60	(1.90)/(4.67)	-56.7
	Moderate	Very low	3.97	(1.83)/(4.32)	-26.5	6.27	(1.56)/(3.79)	+ 4.5

**Table 3** Summary of the regression coefficients, their standard errors (SE)( $\times 10^{-4}$ ) and relative bias of 144,000 simulated datasets on the impact of mixture error model on 2-pollutant Poisson regression by the error variability of PM<sub>2.5</sub> and NO<sub>2</sub>. Results presented for all scenarios ( $N = 3000$  in each row) (Continued)

CRFs <sup>a</sup> :		PM <sub>2.5</sub> : $\beta_1 = 5.4^a$	NO <sub>2</sub> : $\beta_2 = 6^a$
High	Low	4.17 (1.83)/(4.46) -22.7	6.07 (1.59)/(3.70) + 1.2
	Moderate	4.87 (1.81)/(4.28) -9.8	4.87 (1.71)/(4.05) -18.9
	High	5.55 (1.79)/(4.23) + 2.8	2.87 (1.89)/(4.73) -52.1
	Very low	3.03 (1.87)/(4.45) -44.0	6.61 (1.54)/(3.66) + 10.1
	Low	3.30 (1.86)/(4.51) -38.9	6.38 (1.57)/(3.63) + 6.4
	Moderate	4.07 (1.85)/(4.42) -24.7	5.07 (1.70)/(4.12) -15.5
	High	4.54 (1.84)/(4.41) -15.8	3.39 (1.88)/(4.60) -43.4

<sup>a</sup> Concentration-response functions for the generation of the health outcome

<sup>b</sup> Moderate error variability as defined in Table 1. Very low = 0.1 x Moderate, Low = 0.5 x Moderate, High = 1.3 x Moderate

<sup>c</sup> SE<sub>W</sub>: Within-simulations (or model-based) standard error, SE<sub>B</sub>: Between-simulations (or empirical) standard error

<sup>d</sup> Relative bias =  $\frac{\hat{\beta} - \beta}{\beta}$

(Classical, Berkson) percentages: (43,57%) for PM<sub>2.5</sub>, (33,67%) for NO<sub>2</sub>

concentrations) and the exposure metric (e.g. 24-h mean/max, ambient or personal exposure) may vary by study area. When various values for the correlation between the exposures were assessed, mortality estimates did not change significantly. This is important because correlation between pollutant concentrations is the most common factor discussed in air pollution epidemiological studies as driving unreliability in multi-pollutant models. Day et al. (2004) have reached to similar conclusions for the correlations between the exposures and between the errors except for the case of high correlations (both > 0.8), for a nutritional epidemiology study context

[35]. No change was observed in the relative bias with different “true” CRFs, while a small, false positive effect was observed for the pollutant assumed to have no true effect, when both exposures are measured with error. In contrast, when we assumed multiplicative instead of additive error (i.e. a less likely scenario), both pollutant health effect estimates were > 85% biased. Finally, only the PM effect estimate underestimation increased when varying the classical-Berkson ratio of the mixture.

Effect transfer was clearly observed concluding that less precise measurements for one pollutant yield more bias while the co-pollutant effect estimates were closer

**Table 4** Summary of the regression coefficients, their standard errors (SE)( $\times 10^{-4}$ ) and the percentage decrease from single- to multi-pollutant model estimates for 144,000 simulated datasets on the impact of three error models (classical, Berkson and mixture) on 2-pollutant Poisson regression. Results presented for all scenarios ( $N = 144,000$  in each row)

Exposure Model	$\hat{\beta}_1$	(SE <sub>W</sub> )/(SE <sub>B</sub> ) <sup>a</sup>	Bias (%) <sup>b</sup>	Change (%) <sup>c</sup>	$\hat{\beta}_2$	(SE <sub>W</sub> )/(SE <sub>B</sub> ) <sup>a</sup>	Bias (%) <sup>b</sup>	Change (%) <sup>c</sup>
True:								
Multi-Pollutant	5.40	(1.77)/(4.18)	-	+ 32.6	5.99	(1.55)/(3.66)	-	+ 20.5
Single-Pollutant	7.16	(1.71)/(4.06)	-		7.22	(1.50)/(3.55)	-	
Classical:								
Multi-Pollutant	4.65	(1.57)/(3.84)	-13.8	+ 24.7	4.77	(1.35)/(3.40)	-20.5	+ 17.2
Single-Pollutant	5.80	(1.54)/(3.80)	+ 7.3		5.59	(1.32)/(3.38)	-6.9	
Berkson:								
Multi-Pollutant	5.75	(2.41)/(6.32)	+ 6.4	+ 24.7	5.93	(2.17)/(5.49)	-1.2	+ 18.0
Single-Pollutant	7.17	(2.36)/(6.10)	+ 32.8		7.00	(2.13)/(5.37)	+ 16.7	
Mixture:								
Multi-Pollutant	5.03	(1.80)/(4.37)	-6.9	+ 22.5	4.97	(1.68)/(4.25)	-17.1	+ 17.7
Single-Pollutant	6.16	(1.76)/(4.28)	+ 14.1		5.85	(1.62)/(4.21)	-2.4	

<sup>a</sup> SE<sub>W</sub>: Within-simulations (or model-based) standard error, SE<sub>B</sub>: Between-simulations (or empirical) standard error

<sup>b</sup> Relative bias =  $\frac{\hat{\beta} - \beta}{\beta}$

<sup>c</sup> Percentage change from multi- to single-pollutant estimate =  $\frac{\hat{\beta}_s - \hat{\beta}_m}{\hat{\beta}_m}$

(Classical, Berkson) percentages: (43,57%) for PM<sub>2.5</sub>, (33,67%) for NO<sub>2</sub>

**Table 5** Summary of the regression coefficients, their standard errors (SE)( $\times 10^{-4}$ ) and relative bias of 48,000 simulated datasets on the impact of mixture error model on 2-pollutant Poisson regression. Results presented for the core scenario (Area: Europe, Error type: Additive-Mixture) and sensitivity analyses (N = 48,000 in each row)

Sensitivity Analysis	CRFs <sup>a</sup> : Scenario	PM <sub>2.5</sub> : $\beta_1 = 5.4^a$			NO <sub>2</sub> : $\beta_2 = 6^a$		
		$\hat{\beta}_1$	(SE <sub>W</sub> )/(SE <sub>B</sub> ) <sup>b</sup>	Bias (%) <sup>c</sup>	$\hat{\beta}_2$	(SE <sub>W</sub> )/(SE <sub>B</sub> ) <sup>b</sup>	Bias (%) <sup>c</sup>
Main Analysis (Europe-Mixture)		5.33	(1.49)/(3.57)	-1.3	4.36	(2.00)/(5.10)	-27.4
Different "true" CRFs	Low effect CRF <sup>d</sup>	2.66	(1.50)/(3.53)	-1.5	2.19	(2.00)/(4.79)	-27.0
	High effect CRF <sup>e</sup>	10.66	(1.47)/(3.73)	-1.3	8.78	(1.97)/(6.05)	-26.9
	Only PM <sub>2.5</sub> effect	4.85	(1.50)/(3.57)	-10.2	0.13	(2.02)/(4.76)	-
	Only NO <sub>2</sub> effect	0.44	(1.50)/(3.55)	-	4.35	(2.02)/(5.05)	-27.5
Mixture error percentages	(Classical,Berkson) PM <sub>2.5</sub> : (55,45%), NO <sub>2</sub> : (45,55%)	5.20	(1.46)/(3.54)	-3.6	4.47	(1.83)/(4.59)	-25.6
	(Classical,Berkson) PM <sub>2.5</sub> : (70,30%), NO <sub>2</sub> : (60,40%)	5.08	(1.42)/(3.45)	-6.0	4.48	(1.70)/(4.26)	-25.4
Error type	Multiplicative	0.83	(0.27)/(1.92)	-84.5	0.61	(0.27)/(1.68)	-90.0

<sup>a</sup>Concentration-response functions for the generation of the health outcome

<sup>b</sup> SE<sub>W</sub>: Within-simulations (or model-based) standard error, SE<sub>B</sub>: Between-simulations (or empirical) standard error

<sup>c</sup> Relative bias =  $\frac{(\hat{\beta} - \beta)}{\beta}$

<sup>d</sup> Half the CRF from Mills et al. 2006

<sup>e</sup> Twice the CRF from Mills et al. 2006

to the true. This decrease in the bias of the co-pollutant, however, can be regarded as due to the net effect of underestimation due to ME and overestimation due to effect transfer; the latter cancelling out the effect of the former. Szpiro et al. (2011) showed in their simulation study that more accurate exposure predictions do not necessarily improve the health effect estimates [22]. They considered the effects of long-term exposure to air pollution and on comparisons between exposures from correctly specified and misspecified prediction models. However, similar approaches in a multi-pollutant framework have shown that measurement error bias can be severe and correcting for it can strengthen the exposure-response associations [36]. Time-series studies showed that health effect estimates from modelled data are more prone to ME than from measured concentrations [37]. Goldman et al. (2011) reported that spatial error, (only a part of our error decomposition), attenuated the risk ratios from 19 to 31% for primary pollutants (including NO<sub>2</sub>), but only from 2 to 9% for secondary pollutants (PM<sub>2.5</sub> regarded as such) [30]. These values are close to our overall bias estimates of 17 and 7% respectively, even though their characterisation of NO<sub>2</sub> and PM<sub>2.5</sub> as primary and secondary pollutants respectively might be questionable. Similarly, Dionisio et al. (2016) fitting two-pollutant time-series models with additive and multiplicative error reported total effect attenuation up to 85% for NO<sub>2</sub> (close to our estimates for multiplicative error), indicating multi-pollutant model estimates are even more susceptible to ME [38]. Blangiardo et al. (2019) also found, under a Bayesian framework, that NO<sub>2</sub> effects were considerably biased when error-prone

concentrations were used [39]. However, they focused on collinearity in multi-pollutant models without assessing error structures/types.

When misspecified single-pollutant models were fitted, increased effect estimates were observed compared to two-pollutant model coefficients. Interestingly, with mixture error-prone variables, the sum of the single-pollutant model effect estimates was closer to the true combined effect compared with the corresponding sum from multi-pollutant models (due to an accidental cancelling out of overestimation in single-pollutant models from confounding by the co-pollutant and underestimation due to ME). These conclusions hold only when the pollutants are positively correlated. Hence, even if it is accepted that both pollutants have a true causal effect on health [40], should single- or multi-pollutant models be used for quantifying the combined effects of air pollution? Clearly, multi-pollutant models take into account the confounding effects between the pollutants, but if ME is disregarded, biased estimates are produced. This implies that in a "multi-pollutant air quality management framework", ME correction methods should not be neglected [41]. If a correction method cannot be applied, single-pollutant model effects may not be significantly overestimated, but this might only hold under some conditions (e.g. specific correlations between pollutants and/or between errors).

Several previous studies have considered the effects of a mixture of classical and Berkson error in exposures other than air pollution. Mallick et al. (2002) found that the mixture error bias in their relative risks for thyroid disease and radiation fallout ranged from 3.2 to 42.7% [15]. Tapsoba et al. (2019) studying medications in HIV

patients report biases from 0 to 22% depending on the correction method [33]. These values are close to our findings, as is their assumed percentage of Berkson error in the exposure that lies between 20 and 80%. In contrast, Deffner et al. (2018), examining the effects of ultra-fine particles on heart rate, reported that mixture error had little impact on their results [17]. This, however, may be due to their error definitions: they assumed that total personal measurements include only classical error, measurements from fixed sites only Berkson.

This simulation study has some limitations. Firstly, several of our assumptions increased the uncertainty of our estimates. For example, we used pooled estimates for the error structures across wide study areas (e.g. cities within our areas may have different concentrations), and approximations for the correlation between exposures and between errors. However, to the best of our knowledge, no study has attempted to quantify the important error variables in specific areas, apart from Dionisio et al. (2014) in Atlanta (using a different, more spatial framework than our study) [23]. We used more generalisable inputs provided from previously published work [24]. In addition, equal exposure misclassification across days was assumed, increasing the uncertainty about the relationship between sources of spatial and temporal variations. We could not investigate the dependency or correlation between daily exposures and how this might interact with the spatial variation. Moreover, we assumed a constant percent error model additive on the logarithmic scale which implies that at low exposures, error is also low, and vice versa. Error may be higher at high exposures, but it might not drop substantially at high concentrations. In future work, we plan to use real data from panel studies to identify the gaps in our information and describe properly the error structures of the air pollutants.

This study discusses  $PM_{2.5}$  and  $NO_2$  and how to separate their effects on health. There is debate over whether  $NO_2$  is acting as an indicator for other traffic pollutants. To the extent that these other pollutants have similar spatial error characteristics and some, e.g. CO, ultrafine particles, may also have greater infiltration indoors than  $PM_{2.5}$ , our findings regarding effect transfer in multi-pollutant models with  $PM_{2.5}$  will also apply. Moreover, our work could not quantify the error structures of ozone so could not assess its exposure misclassification bias. This work was on two pollutants that are generally positively correlated. Ozone, due to its formation and characteristics, can be negatively correlated with  $PM_{2.5}$  and  $NO_2$ , and this may change our conclusions. However, the correlations are not expected to be very high, and according to this and previous studies, the effect of pollutant correlations on ME bias is not expected to be substantial [35]. We are currently working on this, using

previously analysed raw data from a panel study on schoolchildren [42]. We will also assess the impact of ME on the shape of the exposure-response curve. This work did not find any significant distortions for the shape of the CRFs for the health effects of short-term exposures when the error is additive, but the effects of multiplicative error might be more profound especially for the identification of the long-term effects of air pollution, for which the CRFs are used in cost-benefit analyses [43]. We addressed effect estimates from short-term exposures. Future work should address the multi-pollutant model estimates for long-term exposures which may be more biased, as indicated in studies assessing single-pollutant exposures [31, 32]. Finally, measurement error correction is an important aspect which becomes more complex if we consider expanding the findings from exposure to air pollution to the whole exposome, i.e. the totality of environmental exposures throughout a lifetime. In such studies, variable selection methods are usually used, which rely on empirical data and may create false positive and false negative selections in the presence of ME [44].

In summary, this study quantified the effects of exposure measurement error on multi-pollutant, time-series model estimates. Using simulations, under an extensive range of scenarios, we showed that non-trivial underestimation in health effect estimates can result from measurement error, especially for  $NO_2$ , which was found to be more prone to error, but for  $PM_{2.5}$  as well. We recommend that ME should be considered in every epidemiological analysis assessing exposures prone to large ME, and that studies of personal exposure should provide information on relevant error parameters, such as correlation between errors and error variability, in order to better understand the correct error structures of the pollutants. It is important that correct health effect estimates should be derived in order, not only to separate the independent effects of air pollutants, but also to correctly quantify the health impacts of air pollution, inform interpretation and recommend future approaches for policy making.

### Supplementary information

The online version contains supplementary material available at <https://doi.org/10.1186/s12940-021-00757-4>.

#### Additional file 1.

### Acknowledgements

We thank Professor Ruth Keogh (London School of Hygiene and Tropical Medicine) for her useful feedback in our analysis. We also thank our colleagues from Environmental Research Group (Imperial College London), Dr. Ben Barratt and Dr. David Green, for providing data used as simulation inputs.

**Authors' contributions**

DE: Statistical Methodology, Statistical Analysis, Writing - Original Draft, Review & Editing, KK: Conceptualisation, Statistical Methodology, Writing - Review & Editing, JS: Methodology, Writing - Review & Editing, HW: Conceptualisation, Statistical Methodology, Writing - Review & Editing, Supervision. All authors read and approved the final manuscript.

**Funding**

This work was supported by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Health Impact of Environmental Hazards at King's College London in partnership with Public Health England (PHE) and Imperial College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health & Social Care or Public Health England.

**Availability of data and materials**

The datasets used and analysed during the current study are available from the corresponding author on reasonable request. This work was produced using simulated data.

**Declarations****Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

<sup>1</sup>Environmental Research Group, School of Public Health, Imperial College London, Michael Uren Biomedical Engineering Hub, White City Campus, Wood Lane, W12 0BZ, London, UK. <sup>2</sup>NIHR HPRU in Environmental Exposures and Health, Imperial College London, London, UK. <sup>3</sup>National and Kapodistrian University of Athens, Medical School, Athens, Greece. <sup>4</sup>Department of Environmental Health, T.H. Chan School of Public Health, Harvard University, Boston, MA, USA.

Received: 16 March 2021 Accepted: 7 June 2021

Published online: 24 August 2021

**References**

- Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K, et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the global burden of diseases study 2015. *Lancet*. 2017;389(10082):1907–18. [https://doi.org/10.1016/S0140-6736\(17\)30505-6](https://doi.org/10.1016/S0140-6736(17)30505-6).
- World Health Organization (WHO). Review of evidence on health aspects of air pollution—REVIHAAP project: final technical report. Bonn: WHO European Centre for Environment and Health; 2013.
- Health Effects Institute (HEI). State of Global Air 2019.
- Kelly FJ, Fussell JC. Air pollution and public health: emerging hazards and improved understanding of risk. *Environ Geochem Health*. 2015;37(4):631–49. <https://doi.org/10.1007/s10653-015-9720-1>.
- Dominici F, Peng RD, Barr CD, Bell ML. Protecting human health from air pollution: shifting from a single-pollutant to a multi-pollutant approach. *Epidemiology*. 2010;21(2):187.
- Stafoggia M, Breitner S, Hampel R, Basagaña X. Statistical approaches to address multi-pollutant mixtures and multiple exposures: the state of the science. *Current environmental health reports*. 2017;4(4):481–90. <https://doi.org/10.1007/s40572-017-0162-z>.
- Davalos AD, Luben TJ, Herring AH, Sacks JD. Current approaches used in epidemiologic studies to examine short-term multipollutant air pollution exposures. *Ann Epidemiol*. 2017;27(2):145–53.e1.
- Zidek JV, Wong H, Le ND, Burnett R. Causality, measurement error and multicollinearity in epidemiology. *Environmetrics*. 1996;7(4):441–51. [https://doi.org/10.1002/\(SICI\)1099-095X\(199607\)7:4<441::AID-ENV2263.0.CO;2-V](https://doi.org/10.1002/(SICI)1099-095X(199607)7:4<441::AID-ENV2263.0.CO;2-V).
- Zeger SL, Thomas D, Dominici F, Samet JM, Schwartz J, Dockery D, et al. Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environ Health Perspect*. 2000;108(5):419–26. <https://doi.org/10.1289/ehp.00108419>.
- van Smeden M, Lash TL, Groenwold RH. Reflection on modern methods: five myths about measurement error in epidemiological research. *Int J Epidemiol*. 2020;49(1):338–47. <https://doi.org/10.1093/ije/dy251>.
- Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. Measurement error in nonlinear models: a modern perspective: CRC press; 2006. <https://doi.org/10.1201/9781420010138>.
- Samoli E, Butland BK. Incorporating measurement error from modeled air pollution exposures into epidemiological analyses. *Current environmental health reports*. 2017;4(4):472–80. <https://doi.org/10.1007/s40572-017-0160-1>.
- Shaw PA, Deffner V, Keogh RH, Tooze JA, Dodd KW, Küchenhoff H, et al. Epidemiologic analyses with error-prone exposures: review of current practice and recommendations. *Ann Epidemiol*. 2018;28(11):821–8. <https://doi.org/10.1016/j.annepidem.2018.09.001>.
- Van Roosbroeck S, Li R, Hoek G, Lebret E, Brunekreef B, Spiegelman D. Traffic-related outdoor air pollution and respiratory symptoms in children: the impact of adjustment for exposure measurement error. *Epidemiology*. 2008;19(3):409–16. <https://doi.org/10.1097/EDE.0b013e3181673bab>.
- Mallick B, Hoffman FO, Carroll RJ. Semiparametric regression modeling with mixtures of Berkson and classical error, with application to fallout from the Nevada test site. *Biometrics*. 2002;58(1):13–20. <https://doi.org/10.1111/j.0006-341X.2002.00013.x>.
- Heid I, Küchenhoff H, Miles J, Kreienbrock L, Wichmann H. Two dimensions of measurement error: classical and Berkson error in residential radon exposure assessment. *J Expo Sci Environ Epidemiol*. 2004;14(5):365–77. <https://doi.org/10.1038/sj.jea.7500332>.
- Deffner V, Küchenhoff H, Breitner S, Schneider A, Cyrus J, Peters A. Mixtures of Berkson and classical covariate measurement error in the linear mixed model: Bias analysis and application to a study on ultrafine particles. *Biom J*. 2018;60(3):480–97. <https://doi.org/10.1002/bimj.201600188>.
- Sarnat SE, Klein M, Sarnat JA, Flanders WD, Waller LA, Mulholland JA, et al. An examination of exposure measurement error from air pollutant spatial variability in time-series studies. *J Expo Sci Environ Epidemiol*. 2010;20(2):135–46. <https://doi.org/10.1038/jes.2009.10>.
- Setton E, Marshall JD, Brauer M, Lundquist KR, Hystad P, Keller P, et al. The impact of daily mobility on exposure to traffic-related air pollution and health effect estimates. *J Expo Sci Environ Epidemiol*. 2011;21(1):42–8. <https://doi.org/10.1038/jes.2010.14>.
- Strickland MJ, Gass KM, Goldman GT, Mulholland JA. Effects of ambient air pollution measurement error on health effect estimates in time-series studies: a simulation-based analysis. *J Expo Sci Environ Epidemiol*. 2015;25(2):160–6. <https://doi.org/10.1038/jes.2013.16>.
- Gryparis A, Paciorek CJ, Zeka A, Schwartz J, Coull BA. Measurement error caused by spatial misalignment in environmental epidemiology. *Biostatistics*. 2009;10(2):258–74. <https://doi.org/10.1093/biostatistics/kxn033>.
- Szpiro AA, Paciorek CJ, Sheppard L. Does more accurate exposure prediction necessarily improve health effect estimates? *Epidemiology*. 2011;22(5):680–5. <https://doi.org/10.1097/EDE.0b013e3182254cc6>.
- Dionisio KL, Baxter LK, Chang HH. An empirical assessment of exposure measurement error and effect attenuation in bipollutant epidemiologic models. *Environ Health Perspect*. 2014;122(11):1216–24. <https://doi.org/10.1289/ehp.1307772>.
- 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.
- Katsouyanni K, Samet JM, Anderson H, et al. Air pollution and health: a European and North American approach (APHENA). *Res Rep Health Eff Inst*. 2009;142:5–90.
- Sarnat JA, Schwartz J, Catalano PJ, Suh HH. Gaseous pollutants in particulate matter epidemiology: confounders or surrogates? *Environ Health Perspect*. 2001;109(10):1053–61. <https://doi.org/10.1289/ehp.011091053>.
- Sarnat JA, Brown KW, Schwartz J, Coull BA, Koutrakis P. Ambient gas concentrations and personal particulate matter exposures - implications for studying the health effects of particles. *Epidemiology*. 2005;16(3):385–95. <https://doi.org/10.1097/01.ede.0000155505.04775.33>.
- Sarnat SE, Coull BA, Schwartz J, Gold DR, Suh HH. Factors affecting the association between ambient concentrations and personal exposures to particles and gases. *Environ Health Perspect*. 2006;114(5):649–54. <https://doi.org/10.1289/ehp.8422>.

29. Armstrong BG. Effect of measurement error on epidemiological studies of environmental and occupational exposures. *Occup Environ Med*. 1998; 55(10):651–6. <https://doi.org/10.1136/oem.55.10.651>.
30. Goldman GT, Mulholland JA, Russell AG, et al. Impact of exposure measurement error in air pollution epidemiology: effect of error type in time-series studies. *Environ Health*. 2011;10(61):10.1186.
31. Butland BK, Samoli E, Atkinson RW, Barratt B, Beevers SD, Kitwiroon N, et al. Comparing the performance of air pollution models for nitrogen dioxide and ozone in the context of a multilevel epidemiological analysis. *Environmental Epidemiology*. 2020;4(3):e093. <https://doi.org/10.1097/EE9.000000000000093>.
32. Samoli E, Butland BK, Rodopoulou S, Atkinson RW, Barratt B, Beevers SD, et al. The impact of measurement error in modeled ambient particles exposures on health effect estimates in multilevel analysis: a simulation study. *Environmental Epidemiology*. 2020;4(3):e094. <https://doi.org/10.1097/EE9.000000000000094>.
33. Tapsoba JD, Chao EC, Wang C-Y. Simulation extrapolation method for Cox regression model with a mixture of Berkson and classical errors in the covariates using calibration data. *The international journal of biostatistics*. 2019;15(2). <https://doi.org/10.1515/ijb-2018-0028>.
34. Mills I, Atkinson R, Anderson H, Maynard R, Strachan D. Distinguishing the associations between daily mortality and hospital admissions and nitrogen dioxide from those of particulate matter: a systematic review and meta-analysis. *BMJ Open*. 2016;6(7):e010751. <https://doi.org/10.1136/bmjopen-2015-010751>.
35. Day NE, Wong MY, Bingham S, Khaw KT, Luben R, Michels KB, et al. Correlated measurement error—implications for nutritional epidemiology. *Int J Epidemiol*. 2004 Dec 1;33(6):1373–81. <https://doi.org/10.1093/ije/dyh138>.
36. Bergen S, Sheppard L, Kaufman JD, Szpiro AA. Multipollutant measurement error in air pollution epidemiology studies arising from predicting exposures with penalized regression splines. *Journal of the Royal Statistical Society. Series C, Applied statistics*. 2016 Nov;65(5):731.
37. Butland BK, Armstrong B, Atkinson RW, Wilkinson P, Heal MR, Doherty RM, et al. Measurement error in time-series analysis: a simulation study comparing modelled and monitored data. *BMC Med Res Methodol*. 2013; 13(1):136. <https://doi.org/10.1186/1471-2288-13-136>.
38. Dionisio KL, Chang HH, Baxter LK. A simulation study to quantify the impacts of exposure measurement error on air pollution health risk estimates in copollutant time-series models. *Environ Health*. 2016;15(1):14. <https://doi.org/10.1186/s12940-016-0186-0>.
39. Blangiardo M, Pirani M, Kanařka L, Hansell A, Fuller G. A hierarchical modelling approach to assess multi pollutant effects in time-series studies. *PLoS One*. 2019 Mar 4;14(3):e0212565. <https://doi.org/10.1371/journal.pone.0212565>.
40. Schwartz J, Fong K, Zanobetti A. A national multicity analysis of the causal effect of local pollution, NO<sub>2</sub>, and PM<sub>2.5</sub> on mortality. *Environ Health Perspect*. 2018;126(8):087004.
41. Mauderly JL, Burnett RT, Castillejos M, Özkaynak H, Samet JM, Stieb DM, et al. Is the air pollution health research community prepared to support a multipollutant air quality management framework? *Inhal Toxicol*. 2010; 22(sup1):1–19. <https://doi.org/10.3109/08958371003793846>.
42. Samoli E, Dimakopoulou K, Evangelopoulos D, Rodopoulou S, Karakatsani A, Veneti L, et al. Is daily exposure to ozone associated with respiratory morbidity and lung function in a representative sample of schoolchildren? Results from a panel study in Greece. *J Expo Sci Environ Epidemiol*. 2017; 27(3):346–51. <https://doi.org/10.1038/jes.2016.32>.
43. Crump KS. The effect of random error in exposure measurement upon the shape of the exposure response. *Dose-Response*. 2005 Oct 1;3:4.
44. Agjer L, Slama R, Basagaña X. Relying on repeated biospecimens to reduce the effects of classical-type exposure measurement error in studies linking the exposome to health. *Environ Res*. 2020 Jul 1;186:109492. <https://doi.org/10.1016/j.envres.2020.109492>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

