**ONLINE SUPPLEMENT**

**Exposure assessment – Air Pollution Mixture Levels and Temperature**

To assign robust residential address level exposures to elemental carbon (EC), organic carbon (OC), nitrate (NO3−), sulfate (SO42−), and ammonium (NH4+), we used a novel and well validated hybrid model.1 The model incorporates a hybrid approach adding GEOS-Chem (a chemical transport model) which allows us to include time-space continuous estimates of atmospheric pollutants of PM2.5 and its speciated components. We calibrate these daily GEOS-Chem data using EPA monitors and additional spatio-temporal predictors such as meteorological variables, land-use terms and spatial-temporal lagged terms. We use backward propagation neural network approaches to calibrate then GEOS-Chem predictions from 2001 to 2010 for the Northeastern United States at 1 km × 1 km resolution. The model exhibited excellent performance with ten-fold cross-validation R2 for PM2.5 mass (all data R2 0.85, yearly values: 0.80–0.88) and PM2.5 components (R2 for individual components of 0.70–0.80).

For NO2 residential address level exposures, we used an ensemble approach.2 We used multiple machine learning models each running discretely and then ensembled the predictions, allowing us to fuse these estimations for better overall performance. The ML models we trained included neural networks, random forest, and gradient boosting. We used additional spatio-temporal predictors such as meteorological variables, land-use terms and CTM. This NO2 model covers the entire contiguous U.S. with daily predictions on 1-km-level grid cells from 2000 to 2016. The ensemble produced a cross-validated R2 of 0.788 overall, a spatial R2 of 0.844, and a temporal R2 of 0.729.

We used a very similar approach to predict daily O3 exposures again at a residential level.3 Here we used a neural network adding convolutional layers, which use convolution kernels to aggregate nearby information to account for spatial and temporal autocorrelation. We trained the model against EPA daily maximum ozone data across the continental USA from 2000 to 2012. Cross-validated R2 ranged from 0.74 to 0.80 (mean 0.76) for predictions on our 1 km×1 km grid cells, showing good model performance.

Finally, for daily air temperature (Ta) residence level estimates we used a geostatistical approach incorporating linear mixed effect models, inverse distance weighted interpolations, and thin plate splines (using a smooth nonparametric function of longitude and latitude). To first calibrate surface temperature (Ts) and Ta measurements, we regressed Ta measurements against day-specific random intercepts, and fixed and random Ts slopes. Then to capture the ability of neighboring cells to fill in the cells with missing Ts values, we regressed the Ta predicted from the first mixed effects model against the mean of the Ta measurements on that day, separately for each grid cell. Out-of-sample tenfold cross-validation was used to quantify the accuracy of our predictions. Our model performance was excellent for both days with available Ts and days without Ts observations (mean out-of-sample R2 = 0.95 and R2 = 0.94 respectively). 4

**Statistical Analysis – BKMR-DLM**

Interest focuses on estimating the association between repeated measures of exposure to M pollutants assessed at gestational weeks and a lung function variable *Y* with two primary objectives: 1) to identify critical windows during which the exposure to each component is associated with child’s lung function and 2) estimate the exposure-response relationship while allowing for a nonlinear and non-additive relationship among the multiple components in the mixture and lung function.

We employed a newly developed Bayesian Kernel Machine Regression Distributed Lag Model (BKMR-DLM) detailed elsewhere.5 BKMR-DLM is a form of kernel machine regression that estimates weight functions that up- or down-weight exposure time periods, separately for each mixture component, during which that component has increased or decreased association with the outcome. Critical windows for mixture components are represented as time periods that are up weighted, while time periods with weight function near zero indict little or no evidence of association between a component and the outcome at that time point. BKMR-DLM simultaneously estimates the component-specific weight functions and potentially nonlinear and non-additive associations between a health outcome and the time-weighted exposures. For simplicity, we present the model using an identity link and normal residuals. The model, which we denote BKMR-DLM, is

where and is a smooth function representing the exposure-response surface characterizing the potentially complex association between and .

We use a polynomial kernel function when estimating the exposure-response function h. To parameterize the weight functions, we pre-smooth the exposures with a penalized spline and then use the smooth eigenfunctions of each component as basis functions to estimate smooth weight functions for that component. We use the eigenfunctions that explain 95% of the variation in the smoothed exposures. For model identifiability, the weight functions for times =[1,…,T] and component are constrained so that and . Because of this constraint, the shape of the exposure-response function determines the direction of the exposure-response functions, while the magnitude of the weight function determines which time periods contribute to the exposure response relationship. The sign of the weight function does not impact the direction of the exposure-response function.

We applied BKMR-DLM to lung-function outcomes in the ACCESS cohort to obtain a data-driven weight of each mixture component *m* at each time point *t*. We extracted the estimated temporally-weighted exposure, which are the inner product (i.e. weighted sum) of the estimated weight function and repeated measures of exposure. This results in a BKMR-DLM model-weighted prenatal exposure mixture profile (i.e. a time-weighted average of each pollutant), for each participant. To examine whether the exposure-response pattern was linear or non-linear, we inspected the BKMR-DLM estimated associations, and also applied BKMR to the estimated time-weighted exposures as a check on these estimates.6 We found uniformly linear associations for all exposures and outcomes, thus we performed multivariable linear models using the BKMR-DLM temporal-weighted exposures to examine the association between air pollution mixtures and lung functions. All Bayesian models were fit using 100000 Markov chain Monte Carlo iterations. We discarded the first half of iterations for burn-in and kept every 5th iteration to reduce autocorrelation in the posterior samples. We visually evaluated model convergence by inspecting trace plots and comparing multiple model fits with different random seeds. These models were determined to converge, and consistent results were found across models with different random seeds. Models were fitted for FEV1, FVC, FEF25-75, and the FEV1/FVC ratio. In addition to covariates accounted for in spirometry z-score calculations (age, sex, height, and race/ethnicity), we also adjusted for maternal age, education, and smoking status in the BKMR-DLM and the multivariable linear model analyses. Analyses were conducted in *R* (v4.0.2, Vienna, Austria) using the package “regimes”.

(http://anderwilson.github.io/regimes/bkmrdlm.html)

**Supplemental Table S1. Participant Characteristics: ACCESS Study**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|   | **Included in Analysis** |  | **Enrolled Participants** |  |  |
|   | **(N=198)** |   | **(N=483)** |   | **p-value b** |
| **Maternal age at delivery** |  |  |  |  |  |  |  |
|  Age in years (median, IQR a) | 27.2 | (22.6–32.1) |  | 27.5 | (23.2–32.1) |  | 0.42 |
| **Child sex (n, %)** |  |  |  |  |  |  |  |
|  Female | 93 | 47.0 |  | 216 | 44.7 |  | 0.85 |
|  Male | 105 | 53.0 |  | 252 | 52.2 |  |  |
|  Missing | - | - |  | 15 | 3.1 |  |  |
| **Maternal education (n, %)** |  |  |  |  |  |  |  |
|  >12 years | 66 | 33.3 |  | 150 | 31.1 |  | 0.81 |
|  ≤12 years | 132 | 66.7 |  | 287 | 59.4 |  |  |
|  Missing |  - |  - |  | 46 | 9.5 |  |  |
| **Maternal race/ethnicity (n, %)** |  |  |  |  |  |  |  |
|  Non-Hispanic White | 18 | 9.1 |  | 38 | 7.9 |  | 0.57 |
|  Black/Hispanic-Black | 42 | 21.2 |  | 118 | 24.4 |  |  |
|  Non-Black Hispanic | 124 | 62.6 |  | 260 | 53.8 |  |  |
|  Other | 14 | 7.1 |  | 32 | 6.6 |  |  |
|  Unknown/Missing |  - |  - |  | 30 | 7.3 |  |  |
| **a** IQR = interquartile range (25th percentile – 75th percentile). |
| **b** p-values of the tests comparing the available (non-missing) data between those enrolled vs. included in analysis. Wilcoxon ranked sum test for continuous variables, and χ2 test for categorical variables. |

**Figures S1.** Spearman’s correlation coefficients matrices between trimester average levels of air pollution mixtures. Top to bottom is representing first, second, and third trimester specific correlation matrices.

**Figures S2.** Exposure-response function (h) estimated by BKMR. The solid line indicates the main effect estimate, which is the association between a weighted exposure and the outcomes. Each outcome is shown in a panel: (A) FEV1, (B) FVC, (C) FEV1/FVC, and (D) FEF25-75. The shaded ribbon represents the 95% pointwise confidence interval. The model did not suggest non-linearity on the exposure-response function.



**Figures S3.** Total mixture effect’s exposure-response function (h) estimated using BKMR. The solid dots indicate the main effect estimate at different percentile of the exposure level which is the association between a weighted exposure and the outcomes. Each outcome is shown in a panel: (A) FEV1, (B) FVC, (C) FEV1/FVC, and (D) FEF25-75. The error bars represents the 95% pointwise confidence interval. The model did not suggest non-linearity on the exposure-response function.

**Figures** **S4.** Multivariable-adjusted linear regression models predicting children’s spirometry outcomes (A) FEV1, (B) FVC, (C) FEV1/FVC, and (D) FEF25-75, using pregnancy-average prenatal mixture air pollution levels. From left to right, effect estimates of each component in the air pollution mixture including EC, NH4+, NO3−, NO2, O3, OC, and SO42−, are shown. Effect estimates and 95% confidence intervals are plotted in the order of overall sample (all), boys, and girls to elucidate sex-specific associations.



REFERENCES

1. Di Q, Koutrakis P, Schwartz JD. A hybrid prediction model for PM2.5 mass and components using a chemical transport model and land use regression. *Atmospheric Environment* 2016;**131**:390-399.

2. Di Q, Amini H, Shi L, et al. Assessing NO2 Concentration and Model Uncertainty with High Spatiotemporal Resolution across the Contiguous United States Using Ensemble Model Averaging. *Environmental science & technology* 2019.

3. Di Q, Rowland ST, Koutrakis P, Schwartz JD. A hybrid model for spatially and temporally resolved ozone exposures in the continental United States. *Journal of the Air & Waste Management Association* 2017;**67**:39 - 52.

4. Kloog I, Nordio F, Coull B, Schwartz JD. Predicting spatiotemporal mean air temperature using MODIS satellite surface temperature measurements across the Northeastern USA. *Remote Sensing of Environment* 2014;**150**:132-139.

5. Wilson A, Hsu H-HL, Chiu Y-HM, et al. Kernel machine and distributed lag models for assessing windows of susceptibility to environmental mixtures in children’s health studies. *The Annals of Applied Statistics* 2022;**16**(2):1090-1110, 21.

6. Bobb JF, Valeri L, Claus Henn B, et al. Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. *Biostatistics* 2015;**16 3**:493-508.