ehp ENVIRONMENTAL HEALTH PERSPECTIVES
http://www.ehponline.org

Comparing the Health Effects of Ambient Particulate Matter Estimated Using Ground-Based versus Remote Sensing Exposure Estimates

Michael Jerrett, Michelle C. Turner, Bernardo S. Beckerman, C. Arden Pope III, Aaron van Donkelaar, Randall V. Martin, Marc Serre, Dan Crouse, Susan M. Gapstur, Daniel Krewski, W. Ryan Diver, Patricia F. Coogan, George D. Thurston, and Richard T. Burnett

http://dx.doi.org/10.1289/EHP575

Received: 3 August 2015
Revised: 30 June 2016
Accepted: 18 August 2016
Published: 9 September 2016

Environ Health Perspect DOI: 10.1289/EHP575
Advance Publication: Not Copyedited

Comparing the Health Effects of Ambient Particulate Matter Estimated Using Ground-Based versus Remote Sensing Exposure Estimates

Michael Jerrett[1], Michelle C. Turner[2,3,4,5], Bernardo S. Beckerman[15], C. Arden Pope III[6], Aaron van Donkelaar[7], Randall V. Martin[7], Marc Serre[8], Dan Crouse[9], Susan M. Gapstur[10], Daniel Krewski[2,11], W. Ryan Diver[10], Patricia F. Coogan[12], George D. Thurston[13], and Richard T. Burnett[14]

[1]Department of Environmental Health Sciences, Fielding School of Public Health, University of California Los Angeles, Los Angeles, California, United States;

[2]McLaughlin Centre for Population Health Risk Assessment, University of Ottawa, Ottawa, Ontario, Canada;

[3]ISGlobal, Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain;

[4]Universitat Pompeu Fabra (UPF), Barcelona, Spain;

[5]CIBER Epidemiologia y Salud Publica (CIBERESP), Madrid, Spain;

[6]Department of Economics, Brigham Young University, Provo, Utah, United States;

[7]Department of Physics and Atmospheric Science, Dalhousie University, Halifax, Nova Scotia, Canada;

[8]Department of Environmental Sciences and Engineering, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States;

[9]Department of Sociology, New Brunswick Institute of Research, Data and Training, University of New Brunswick, Fredericton, New Brunswick;

[10]Epidemiology Research Program, American Cancer Society, Atlanta, Georgia, United States;

[11]Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Ontario, Canada;

[12]Slone Epidemiology Center, Boston University, Boston, Massachusetts, United States;

[13]New York University School of Medicine, Tuxedo, New York, United States;

[14]Population Studies Division, Health Canada, Ottawa, Ontario, Canada;

[15]Division of Environmental Health Sciences, Public Health Department, University of California at Berkeley, Berkeley, California, United States

Corresponding author: Michael Jerrett, Ph.D.

650 Charles E. Young Drive S, Rm. 56-070 CHS, Mail Code: 177220,

Los Angeles, CA 90095. (310) 825-9037. E-mail: mjerrett@ucla.edu

Short running title: Mortality effects from different exposure models

Acknowledgments

Funding was received to support some aspects of this work from CDC Award 200-2010-37394 A Multilevel Geographic Model for Environmental Public Health Tracking, the National Institute of Environmental Health Science 5R01ES019573-04 Air Pollution and Risk of Incident Hypertension and Diabetes in U.S. Black Women, and from a research contract granted by Health Canada. Michelle C. Turner was funded by a Government of Canada Banting Postdoctoral Fellowship. George D. Thurston was supported by the NYU-NIEHS Center of Excellence, (5P30ES000260).

Competing financial interests declaration: Daniel Krewski is the Chief Executive Officer and Chief Risk Scientist of Risk Sciences International (RSI). RSI is a consulting company specializing in the assessment, management, and communication of health and environmental risks and their broader impacts on both public and private interests. RSI represents a wide range of private and public sector clients. Dr. Krewski's involvement in this company has not, however, influenced the analysis, results or interpretation reported in the above paper. Dr. Krewski also hold an NSERC Industrial Research Chair, which is half funded by industrial interests. The Chair is administered by SNERC independently and the industrial funders have no influence on the research undertaken by Dr. Krewski. The other authors declare they have no actual or potential competing financial interests.

ABSTRACT

Background: Remote sensing (RS) is increasingly used for exposure assessment in epidemiological and burden of disease studies, including those investigating whether chronic exposure to ambient fine particulate matter ($PM_{2.5}$) is associated with mortality.

Objectives: To compare relative risk estimates of mortality from diseases of the circulatory system for $PM_{2.5}$ modeled from RS with that for $PM_{2.5}$ modeled using ground-level information.

Methods: We geocoded the baseline residence of 668,629 American Cancer Society Cancer Prevention Study II (CPS-II) cohort participants followed from 1982 to 2004 and assigned $PM_{2.5}$ levels to all participants using seven different exposure models. Most of the exposure models were averaged for the years 2002-2004, while one RS estimate was for a longer, contemporaneous period. We used Cox proportional hazards regression to estimate relative risks (RR) for the association of $PM_{2.5}$ with circulatory mortality and ischemic heart disease.

Results: Estimates of mortality risk differed among exposure models. The smallest relative risk was observed for the RS estimates that excluded ground-based monitors for circulatory deaths (RR = 1.02 (95% confidence interval (CI): 1.00-1.04 per 10 $\mu g/m^3$ increment in $PM_{2.5}$). The largest relative risk was observed for the land use regression model that included traffic information (RR = 1.14, 95% CI: 1.11-1.17 per 10 $\mu g/m^3$ increment in $PM_{2.5}$).

Conclusions: We found significant associations between $PM_{2.5}$ and mortality in every model; however, relative risks estimated from exposure models using ground-based information were generally larger than those estimated with RS alone.

INTRODUCTION

Remote sensing and atmospheric chemistry models play an increasingly important role in exposure assessment for epidemiological and burden of disease studies. A wide array of products produced by several U.S. Federal agencies, such as National Aeronautics and Space Administration (NASA) and the Environmental Protection Agency (EPA), are now available. Sometimes these models form the basis for more complex estimates combining ground-based data or several remote-sensing products.

Several recent epidemiological investigations have used remote sensing for the exposure assessment or as in input into other health impact assessment or variable imputation models. By combining retrievals of aerosol optical depth (AOD) from the Moderate Resolution Imaging SpectroRadiometer (MODIS) and Multi-angle Imaging SpectroRadiometer (MISR) instruments onboard the Terra satellite with the GEOS-Chem model, van Donkelaar, et al., (2010) developed six-year mean global estimates of $PM_{2.5}$ at about 10 km resolution (van Donkelaar et al. 2010). These RS products were designed to avoid reliance on $PM_{2.5}$ monitors, as these RS products can offer information about $PM_{2.5}$ in regions where $PM_{2.5}$ monitors are not generally available, or where there are concerns about $PM_{2.5}$ data quality as, for example, with Tapered Element Oscillating Microbalances (TEOMs). Researchers in Canada have used the van Donkelaar et al. (2010) estimates to assess the health effects of air pollution. Specifically these $PM_{2.5}$ estimates were significantly associated with incidence of diabetes (Chen et al. 2013), diabetes mortality (Brook et al. 2013), and cardiovascular mortality (Crouse et al. 2012; Crouse et al. 2015). These RS estimates have also been used to estimate the global mortality associated with $PM_{2.5}$ (Evans et al. 2013; Lim et al. 2012).

A few studies have attempted to compare systematically the exposure estimates from ground-based versus remote sensing models. Lee et al. (2012) developed national-level models using data from more than 1,300 ground-monitors for $PM_{2.5}$ (Lee et al. 2012). Their results indicated that, within about 98 km of a monitor, the ground-based estimates predicted $PM_{2.5}$ concentrations more accurately than the RS estimates discussed above (van Donkelaar et al. 2010). Beyond 98 km, however, the RS estimates were better predictors of ground-level $PM_{2.5}$. For the most part, the estimates were highly correlated with each other, and the authors concluded that the differences in prediction capacity were fairly small. Another study compared NASA AOD retrievals to ground-based estimates derived from a generalized linear model that included ground information on land use predictors and several statistical smoothing functions. This study concluded that the RS estimates were not generally better predictors than the ground-based models, and after applications of smoothing functions in the models, there was little marginal benefit to the remote-sensing information on predicting ground-level $PM_{2.5}$ (Paciorek and Liu 2009). Subsequent studies have found that ground-based observations can be better predicted using exposure models with RS estimates than without (Beckerman et al. 2013a; Beckerman et al. 2013b; Kloog et al. 2012b; Ma et al. 2014; Vienneau et al. 2013).

RS estimates of air pollution generally lack the fine-scale resolution (<1 km) needed for use in environmental epidemiological studies that aimed at understanding small-area variations in exposure. To achieve horizontal downscaling of the remote sensing estimates, “hybrid” approaches that combine variants of “land use regression models” that predict pollutant concentrations from land use such as road length, traffic density or, open space with RS

measurements are being employed (Beckerman et al. 2013a; Beckerman et al. 2013b; Kloog et al. 2012b; Ma et al. 2014; Vienneau et al. 2013). Through statistical modeling, proxy information about likely locations of pollution at smaller spatial resolution than AOD pixels can essentially distribute the $PM_{2.5}$ estimated from the AOD to its most likely locations within its pixel. These hybrid exposure estimates have been used in a number of epidemiological studies (Jerrett et al. 2013; Kloog et al. 2012a; Kloog et al. 2013; Madrigano et al. 2013).

Although now in broader use, little is known about the impact of using RS estimates on predicted health effects as compared to either monitored data or hybrid models. In this paper we use the American Cancer Society Cancer Prevention Study II (CPS-II) − a well-documented, United States-wide prospective cohort study − to compare various RS, geostatistical and hybrid models in the estimation of circulatory and ischemic heart disease mortality (IHD) associated with ambient $PM_{2.5}$. Recently several papers have been published using 1 km estimates of $PM_{2.5}$ for the United States; some of these more spatially fine-grained estimates use ground data extensively (see, e.g., Lee et al. 2015; van Donkelaar et al. 2015b). Calibration with ground data likely improves the performance in the United States, where there is a large and spatially-wide coverage of ground-based monitors. In other regions that lack extensive monitoring support, such calibration is more challenging. In this paper, therefore, we have included estimates that incorporate ground data and those that rely solely on remote sensing retrievals. This allows us to assess directly the importance of ground data calibration.

METHODS

This section outlines the health data, exposure models, and statistical analyses. Further details are provided in the Supplemental Material.

Health and Demographic Data

In September 1982 and February 1983, volunteers enrolled the CPS-II cohort. In total, 1,184,587 participants at least 30 years of age were enrolled at baseline. The participants were mostly friends and family members of the volunteers. Participants were recruited from all 50 states, the District of Columbia, and Puerto Rico. They completed a four-page, self-administered survey with items on demographic, lifestyle, medical, and other variables, including residential address at baseline. The CPS-II was approved by the Institutional Review Board of Emory University and participants provided informed consent prior to participation. Approval for the analysis in this paper specifically was obtained from the Ottawa Hospital Research Ethics Board and the Committee on the Protection of Human Subjects, University of California, Berkeley.

We geocoded participant residence at baseline, which was then used to assign several exposures, at either the individual participant residence or census tract (CT) of residence (see Pope et al. 2015 for details). After making exclusions for missing residence information and key covariates such as smoking, 668,629 participants remained in the analytical cohort used in this analysis. See Table S1 in the Supplementary Material for a comparison of those included versus those excluded and some commentary on the minor differences between the two groups.

Vital status from 1982 through to 2004 was ascertained using methods documented elsewhere (Jerrett et al. 2009). Briefly, in 1984, 1986, and 1988, vital status was determined by the study volunteers, with subsequent confirmation obtained by the corresponding death certificate. For deaths after 1989, computerized linkage to the National Death Index has been used for follow-up (Calle and Terrell 1993). We focused on mortality from diseases of the circulatory system (ICD9 codes 390-459; ICD10 codes I00-I99) for comparability with existing studies (see, e.g., Hoek et al. 2013). We also examined effects on IHD (ICD 9 codes 410-414; ICD10 codes I20-I25) deaths because this outcome had the largest effect sizes in the ACS cohort (Turner et al. 2016) and thus is amenable to assessing inter-model differences in the exposure assessment. Given evidence that long-term $PM_{2.5}$ exposures may also be associated with diabetic deaths (Brook et al. 2013; Pope et al. 2015), as supplementary analysis, we also examined effects on diabetic deaths (ICD10 code E11), although there are considerably fewer deaths.

Exposure Models

The models are summarized in Table 1 in terms of their spatial and temporal resolution, the types of data used to derive the estimates, and the cross-validation results. First, the remote sensing dataset, mentioned earlier, was developed with the MODIS and MISR satellites with scaling to ground level achieved via a chemical transport model (GEOS–Chem). These initial estimates were produced globally on a 0.1×0.1 degree (~9.8 km) for the years 2001-2006 (van Donkelaar et al. 2010). Additional RS estimates were also included, representing 2002-2004 (van Donkelaar et al., 2015b). These updated estimates were provided at 0.01 x 0.01 (~1 km) degree resolution as produced with an optimal estimation algorithm developed for MODIS observations, and the subsequent inclusion of ground-based observations through a globally applicable (van Donkelaar

et al., 2016) geographically weighted regression that restricted monitors for training to >100 km away. Scaling to years prior to 2004 followed van Donkelaar et al., 2015a, which relied upon trend information (Boys et al. 2014) from the SeaWiFS (Hsu et al. 2013) and MISR satellite instruments. Second, we assigned the Hierarchical Bayesian Model (HBMCMAQ) developed by the U.S. Environmental Protection Agency (McMillan et al. 2010). This model fuses daily estimates from the Community Multi-Scale Air Quality (CMAQ) model with ground observations in a Bayesian modeling regime that essentially upweights the influence of the CMAQ predictions as a function of distance away from the monitor. These estimates were derived nationally for ~36 km grid. We averaged the daily estimates to a 3-year average of 2002-2004 (cf. Turner et al. 2015). Third, we assigned a Bayesian Maximum Entropy (BME) spatiotemporal geostatistical kriging model based on ground observations (~ 9.8 km) (Lee et al. 2012). This model was fit based on 1364 in situ monitors. Fourth, we assigned a hybrid land use regression model using only ground-based inputs where the first stage of the model was fit with a deterministic regression model with monthly pollution as the dependent variable and land use and traffic information as predictors, with the second stage consisting of a BME kriging model of the residuals (BMELUR). Predictions from the two models were combined post hoc to derive the exposure surface, which was averaged over the period 2002-2004. Finally, we developed the fifth model using similar kriging-LUR approach that combined ground-based information with the RS estimates (BMELURRS) (Beckerman et al. 2013b). The variables in final two models were selected with a deletion/substitution/addition algorithm, which relies on v-fold cross-validation to avoid over-fitting to the measured data.

Statistical Models

We employed Cox proportional hazards regression to examine associations of $PM_{2.5}$ exposure with death from diseases of the circulatory system and with IHD while controlling for likely individual and ecological confounders.

We used follow-up time in days from enrollment as the time axis. As in previous analyses (Jerrett et al. 2013), we stratified models by 1-year age categories, sex, and race (white, black, other). This allowed each category to have its own baseline hazard. We included a comprehensive set of individual risk factor variables operationalized in a similar way to those used in previous studies of the CPS-II cohort (Jerrett et al. 2009; Krewski et al. 2009). We used ecological variables extracted from the 1990 and 2000 U.S. Census in the ZIP code neighborhoods of residence to control for potential "contextual" neighborhood effects. We provide details in the Supplemental Material.

As a sensitivity analysis, we estimated selected adjusted hazard ratios using multi-level models that included a random effect term for the county of residence (Jerrett et al. 2009). We also included variables controlling for the size of the metropolitan area of residence, which has been found to influence air pollution-mortality associations (Crouse et al. 2012; Jerrett et al. 2013), and elevation because higher elevations have been related to cardiovascular mortality in this cohort and are generally associated with lower pollution levels (Krewski et al. 2000).

To assess overall model fit, we used the Akaike Information Criterion (AIC). We hypothesized that models with a better overall prediction had lower measurement error and would therefore have lower AIC values and larger coefficients in the Cox regression model.

Our assessments of model fit and effect size might be suggestive about which model provides the best prediction of mortality, but in observational studies we have no way of knowing which model best reflects the true relationship between air pollution and survival. For our main results, therefore, we developed ensemble estimates that pooled the effects from every model into a single estimate. This method derives a weighted average of the coefficients from the various models with the weights defined in terms of the change in the AIC from the model with the minimum AIC (Buckland et al. 1997; Faes et al. 2007). Specifically we computed the weights as follows:

$$w_i = \frac{\exp\left\{-\frac{1}{2}\left(AIC_i - AIC_{mn}\right)\right\}}{\sum_{l=1}^{L}\exp\left\{-\frac{1}{2}\left(AIC_l - AIC_{mn}\right)\right\}}$$

where AIC_i is the Akaike Information Criterion of the i^{th} model and AIC_{mn} denotes the minimum Akaike Information Criterion among the L models examined. Since the number of parameters are identical in all models and since $AIC = -2\log lik + 2k$, $\log lik$ denoting the logarithm of the likelihood function and k the number of parameters in the model, the ensemble weights w_i can be written as $w_i = lik_i \Big/ \sum_{l=1}^{L} lik_l$ where lik_i is the likelihood function for the i^{th} model. In this case

we interpret the ensemble weights as a function of the likelihood and not necessarily the AIC. However, in practice we use the definition of the ensemble weights in terms of the AIC since the value of the likelihood for a study as large as the CPS-II cohort is too big to be calculated with standard computing software.

RESULTS

Descriptive statistics of the analytic cohort are shown in Table S1, which presents the average $PM_{2.5}$ exposures across the strata of the covariates. Few variables appear to be associated with $PM_{2.5}$. African Americans tend to have higher exposures, although they account for a very small proportion of the cohort (3.8%). A slight inverse trend exits for those with higher education having lower pollution levels.

Table 2 presents descriptive statistics for the different exposure models considered. The mean $PM_{2.5}$ estimates are very similar across models. The RS models demonstrate the highest variation, as measured by the standard deviation and the inter-quartile range (IQR). The BMELUR model, however, has the largest overall range.

As shown in Table 3, moderately high to very high correlations exist among $PM_{2.5}$ estimates from the five models (HBMCMAQ, BME, BMELUR, BMELURRS, $PM_{2.5}$ RS GWR CT 02-04) that included ground-based data in various ways, with correlations ranging from $r = 0.71$ to $r = 0.94$. The RS estimates without ground data show similarly high correlations with one another, ranging from $r = 0.84$. The two $PM_{2.5}$ model groupings (ground based versus remote sensing), however, have lower correlations with one another ranging from $r = 0.54$ to 0.67. The one

exception is a moderately high correlation between the two 1 km resolution remote sensing surfaces ($PM_{2.5}$ RS GWR CT 02-04 and $PM_{2.5}$ RS no GWR CT 02-04) that exhibit a moderately high correlation r = 0.78. Models that were assigned at both the participant residence and CT-level had very high correlations (r = 0.94 to 0.99).

Table 4 shows the results from the Cox regression modeling for mortality from diseases of the circulatory system and for IHD. We observe significant associations between particulate matter exposure and death in every model, although substantial variation exists among the magnitudes of the risk estimates. We found minor changes in the estimates when the ecologic confounders were added to the model, and given prior knowledge of how ecologic variables can affect mortality-air pollution associations, we report these as our primary results. For circulatory mortality, we observed the highest relative risks from the BMELUR (RR = 1.14, 95% CI, 1.11-1.17), while the lowest relative risks resulted from the $PM_{2.5}$ RS no GWR CT 02-04 estimate that excluded ground data (RR = 1.02, 95% CI, 1.00-1.04). Relative risks from the other models are closer to those of the BMELUR (RR ~ 1.08-1.12). The ensemble estimate is the same as the BMELUR for circulatory mortality (RR = 1.14, 95% CI, 1.11-1.17). As a sensitivity analysis, we also temporally matched the LURBME to the RS estimate from the 2001-2006 period. The results show slightly larger differences between the two estimates but they are similar in magnitude to those when the temporal periods differed (see Table 4).

Inclusion of the ecological variables had a relatively larger effect on the IHD estimate than with circulatory mortality. Although the inclusion of the ecologic variables diminished the differences in the RRs between the exposure models, the pattern is similar, with $PM_{2.5}$ RS no GWR CT 02-

04 RS producing the lowest RRs. The differences between the other estimates, however, are somewhat less when the ecological covariates are included, with the BME kriging model and the BMELUR having the largest risks, followed closely by BMELURRS and the HBMCMAQ.

We also compared results across the inter-decile range (IDR) of exposure (see Supplemental Material, Table S2), which shows lesser differences between the estimates. We included this analysis to compare the models across the same range of exposure within their own distribution. The relative ordering is maintained for circulatory deaths, with the BMELUR and RS without ground-monitors producing the largest and smallest relative risk estimates, respectively. For IHD deaths, after inclusion of the ecological covariates, many of the estimates are very similar and the RS 01-06 model actually produces slightly larger relative risks (i.e., 1.1 versus 1.09), although the BMELUR model still has the lowest AIC of all the models indicating that this model has the best model fit.

We also restricted the analysis to only those subjects who resided in cities with government monitoring stations (Table 5). This allowed us to compare the seven exposure models with those using only the spatial average per county, similar to earlier reports from the cohort that used only the central monitoring data (Jerrett et al. 2009; Pope et al. 2002). Here we see similar ordering; however, the RS model estimates tend to be even lower than before for this subset of the cohort. For circulatory deaths, the $PM_{2.5}$ RS GWR CT 02-04 estimates is smaller than the county-wide average of the government monitor exposure estimates, which effectively have only one estimate per county. For IHD, all remote sensing estimates are smaller than the county-wide estimate.

Results indicate that models with a lower AIC (i.e. better model fit) generally have a higher relative risk estimate. For example, the increase in the relative risks of circulatory and IHD deaths with respect to max (AIC)-AIC is clearly seen in Figure S1 (R^2=0.94). The AIC attains its maximum value (corresponding to the worst fitting model) when using the RS-based exposure assessment methods. The AIC has the lowest levels for the model including ground-based exposure methods, indicating improved model fit. Of these, the HBMCMAQ method has a larger AIC (poorer fit) than any of the BME methods, with BMELUR resulting in the smallest AIC (best overall fit) of all the methods. In instances where we used geocoding to the residential address or the CT, when we estimated exposures with both models of exposure assignment, we see slight attenuation of the effects with the CT exposure assignment compared to the residential address. For example, with the BMELUR 02-04, the RR is 1.14 with the residential address, while with the CT assignment it is 1.12. The higher AIC for the CT assignment indicates some degradation in model fit from the CT assignment.

As a sensitivity analysis, we included metropolitan area size as a covariate, given earlier findings that suggested larger cities were associated with both lower mortality and higher pollution. We also included an elevation variable in this analysis. Inclusion of both variables separately or together had little impact on the size or overall pattern of the risks (see Tables S4 and S5 in the Supplemental Material for details). For the two models with the highest and lowest relative risk estimates (the BMELUR and the RS models, respectively), we included a random effect at the county level. With the random effect, we see even larger differences in the size of the relative risk between the BMELUR and RS exposure models than with the standard Cox model. As a final sensitivity analysis, we ran the models using ecological confounders from 2000 instead of

1990 (see Table S6 in the supplementary material). The 2000 ecological variables exert a slightly larger confounding effect on the $PM_{2.5}$ relative risks in all models, but all results remain significantly elevated, and the ordering of the relative risks between models is consistent with what we observed in the earlier analyses using 1990 ecological covariates.

Table S7 shows the results for the diabetes deaths. Here without the ecological covariates only risks from the LURBME models are significantly elevated, while all others include unity in the 95% confidence interval. With the addition of ecological covariates, many of the models do have significantly elevated risks, and the rank ordering among the models follows a similar pattern to what we observed with circulatory and IHD deaths. Specifically the largest risks are observed in models that use the BMELUR with 1.18 (95% CI: 1.05-1.33), while the smallest are in models that use remote sensing with no ground data (RS no GWR CT 02-04) with RR = 1.01 (95% CI: 0.92-1.11).

The concentration-response (C-R) curves for the BMELUR and RS models are shown in Figures S2 and S3. We investigated these curves to gain insights into the likely shape of the C-R curves. These curves were based on natural splines with 2 degrees of freedom. As expected from the model results, the BMELUR C-R curve has a steeper slope consistent with the larger coefficient versus the RS effect estimate. The RS model exhibits a declining slope at approximately 15 $\mu g/m^3$.

DISCUSSION

We found statistically significant positive associations between $PM_{2.5}$ exposures and risk of death from circulatory disease and IHD with every exposure model tested for circulatory and IHD deaths. With the smaller number of diabetic deaths, we did not observe significant effects in all models for this outcome, but there were significant effects in many of the models after control was applied for ecological confounding. These findings agree with several studies on this cohort, some of which used government monitors (Jerrett et al. 2009; Krewski et al. 2009; Pope et al. 2002), interpolation models (Jerrett et al. 2005), or the hybrid land use regressions that included ground based-information with traffic (Turner et al. 2014; Turner et al. 2015; Pope et al. 2015) and remote sensing with land use (Jerrett et al. 2013). Our current finding strengthens the evidence base for a non-spurious association between $PM_{2.5}$ exposure and mortality because estimates were significant for in most models, regardless of the exposure assessment method.

In general, the findings agree with existing evidence on the associations of $PM_{2.5}$ with CVD outcomes although the estimated associations here are somewhat larger. For example, the ensemble estimate for circulatory deaths was RR =1.14 (95% CI: 1.11, 1.17) or a 14% increase, while a recent meta-analysis estimated a 10.6% (95% CI 5.4, 16.0%) per the same exposure contrast of 10 μ g/m^3 (Hoek et al. 2013). Similarly, a recent analysis of another large, nationwide U.S. cohort found a 10% increase CVD mortality with a RR=1.10, 95% CI=1.05, 1.15 (Thurston et al, 2015). Results for RS models without ground monitors were about one fifth to one half the size of the meta-analysis estimates. Consistent with the Hoek et al. (2013) meta-analysis, we also observed slightly larger associations for IHD than for the broader circulatory

category, although these estimates tended to be attenuated relatively more by addition of ecological covariates than for those in the circulatory mortality category.

Although the finding of significant associations between $PM_{2.5}$ and mortality appears consistent for most specific exposure model tested, the RR estimates varied markedly among models. Compared to past studies, RRs here were larger for ground-based and more sophisticated hybrid models than with central monitors or remote sensing exposure models alone. Among the ground-based exposure models tested, the HBMCMAQ model based on linear geostatistics had the poorest fit and smallest effect size for circulatory and IHD deaths, while the BME models based on Bayesian Epistemic knowledge blending had a better fit and larger effect size. Of the BME exposure models, the best fit and largest association is obtained with BMELUR, a hybrid model containing information on traffic and local land use. On a per microgram basis, the model containing traffic had associations that were more than 2.5 times greater than the remote sensing models for circulatory mortality. This might suggest a higher toxicity for the mixture of $PM_{2.5}$ that originates from traffic or that fine-scale exposure estimates are needed to accurately assess health effects. We have used the 10 μ g/m^3 as our primary comparison because this exposure increment shows the relative difference on a per microgram basis. Even small increments to improve overall accuracy of the exposure can be important for health effects assessment based on where these differences occur spatially. In the case of the BMELUR, the maximal contribution of the traffic variable to the overall model prediction was small, on the order of 1-1.5 μ g/m^3, but the spatial alignment of this to areas with dense traffic appears to capture potentially heightened toxicity from this source or vulnerability of the populations living in areas of high traffic or both, which translates into much larger (2.5 times greater than RS) effects on a

per microgram basis. The differences between the models using the IDR were relatively smaller, which likely indicates that such comparisons are less able to determine essential differences between models that might arise from their ability to detect fine-scale variations near the source. Remote sensing models without ground data (i.e., RS no GWR CT 02-04 and the RS 01-06) also had larger IDRs (i.e. 9.6 $\mu g/m^3$ and 9.9 $\mu g/m^3$, respectively) though the LURBME still had the largest range (26.6 $\mu g/m^3$). Thus, the results also appear sensitive to the relative distributions of various exposure models. On a per microgram basis, however, the relative rank ordering is clear and consistent.

Other emerging fine resolution satellite retrievals (see, e.g., Lyapustin et al. 2011; Lee et al. 2015) may better resolve local aerosol sources, which might align the remote sensing estimates more closely with ground-based models. Remote sensing is being increasingly combined with ground-based and LUR information for overall accuracy and to include finer resolution information (Beckerman et al. 2013; Kloog et al. 2012b; Ma et al. 2014; van Donkelaar et al. 2015b). As these higher resolution models become more widely available, comparing these to other models such as the LURBME will be important for understanding whether, on a per microgram-basis, these hybrid models will detect health risks of similar magnitude to other models. Our initial investigation here with the 1 km resolution RS GWR CT 02-04 suggests that even with ground calibration, these models are not yet detecting risks of similar magnitude to the BMELUR, which included traffic data and smaller-area prediction.

Our findings suggest caution against over-interpreting quantitative estimates of association between ambient $PM_{2.5}$ and mortality based on a specific exposure assessment method. This is

particularly important for estimating the air pollution-related burden of disease, which has relied on a pooling of concentration-response functions from studies with varying exposure assessment methods (Lim et al. 2012). Our work suggests that concentration-response modeling should be based on the most appropriate source of exposure information available. In regions where ground-based monitoring is sparse, the best available option may be to conduct an RS-based exposure assessment and use the best available concentration-response curve (Figure S2). Actual health risks, however, could be even larger than those derived from RS estimates without ground data. Hence, when there is sufficient ground-based monitoring data to calibrate an exposure model, hybrid exposure models should be used. Amongst those, we found that the HBMCMAQ exposure model had a poorer fit than the BME models; amongst those, the BMELUR model based on traffic and local land had the best fit and largest effect size. In that case, our ground-based concentration response curve (Figure S2) can therefore be considered as representing the best predictor of mortality. This conclusion is supported by the ensemble estimate that, due to the superior fit of the BMELUR, ascribed nearly all the weight to this estimate of exposure. Because it reflects a larger effect size, that curve will attribute more deaths attributable to $PM_{2.5}$ than other models, particularly those based on remote sensing with no ground data. Looking forward, an emerging global $PM_{2.5}$ network (SPARTAN) is taking ground-based measurements to address key sources of uncertainty in remote sensing estimates (Snider et al. 2015).

Our evaluation of the concentration-response functions for the BMELUR and RS models represents just two of several possible exposure response functions and is intended purely to add in visualizing whether, with identical spline functions, we observe different shapes of the exposure-response relationship between air pollution and mortality. We caution, however,

against using these visual plots for understanding the underlying exposure-response functions. Such analyses would have to examine some weighted combination of several possible models, rather than the single realization discussed here.

The present study has several strengths. We used a large data set with a long follow up and excellent control for covariates that could confound the air pollution-mortality relationship to estimate mortality associated with particulate air pollution. We also employed a comprehensive suite of exposure models, ranging from those with no ground information such as RS, to chemical transport models fused to ground data with Bayesian methods, to geostatistical kriging models, and finally to hybrid models that included either ground data only combined with advanced interpolation methods or some fusion remote sensing and land use data. This suite of models covers most of the currently available exposure assessment methods likely to be employed in epidemiological analyses of mortality associated with long-term exposure to ambient air pollution.

The study also has several limitations. First, most exposure models were assigned at or near the end of the follow up, largely due to the lack of $PM_{2.5}$ data before 1999. Earlier analyses have shown that the relative spatial pattern likely remains constant in rank ordering over time (Jerrett et al. 2005). Relative declines in $PM_{2.5}$ may have occurred unevenly across the country, potentially resulting in spatial heterogeneity that is not captured by the exposure models and, therefore, differential levels of exposure error in each model. Moreover, study participants may have moved to higher or lower exposure areas, which could again impart error in the risk estimates. The extent to which either of these possible sources of error would influence the effect

estimates from any of the models is difficult to determine. Some models had inherently larger grid areas for exposure assignment than others. Potential error sources in the RS GWR estimate include exclusion of ground-based observations on spatial scales within 100km and scaling to the 2002-2004 period. It would be instructive to revisit these limitations in future work. Similarly, one RS estimate, while contemporaneous with the ground estimates, was of longer duration and ran past the end of the follow up, and this may have introduced additional error that affected the relative size of the estimates. We did, however, conduct a sensitivity analysis where we matched the LURBME to the exact temporal periods of the RS 01-06. The results from this sensitivity analysis were slightly stronger in terms of the difference between the two estimates but were essentially the same as those presented in the main results, suggesting the temporal misalignment had a negligible impact on the overall patterning of the risks. We were unable to quantitatively assess the impact of measurement error with formal models, due to the unavailability of an externally valid "gold standard" to implement a regression calibration model (cf. Molitor et a. 2007).

CONCLUSIONS

We found significant associations between $PM_{2.5}$ exposure estimated using different models and risk of mortality. Relative risks estimated from exposure models using ground-based information were larger than those estimated with only RS.

The range of relative risks seen in this study also suggests new avenues for understanding the health effects of air pollution. This approach would follow the lead of climate models, whereby the various relative risk estimates could be combined or pooled into one estimate that would

capture the range and uncertainty in the estimates. Similar approaches have been used to combine and assemble various estimates of future climate (Intergovernmental Panel on Climate Change 2014), where inherent uncertainty exists and no estimate is objectively superior. As an initial approach, we have developed ensemble estimates. Such approaches could be expanded to supply more accurate estimates of the effects of air pollution exposure on mortality, with appropriate characterizations of model uncertainty.

REFERENCES

Beckerman BS, Jerrett M, Martin RV, van Donkelaar A, Ross Z, Burnett RT. 2013a. Application of the deletion/substitution/addition algorithm to selecting land use regression models for interpolating air pollution measurements in california. Atmos Environ 77:172-177.

Beckerman BS, Jerrett M, Serre M, Martin RV, Lee S-J, van Donkelaar A, et al. 2013b. A hybrid approach to estimating national scale spatiotemporal variability of $PM_{2.5}$ in the contiguous united states. Envir Sci Technol 47:7233-7241.

Boys BL, Martin RV, van Donkelaar A, MacDonell RJ, Hsu NC, Cooper MJ, et al. 2014. Fifteen-year global time series of satellite-derived fine particulate matter. Envir Sci Technol 48:11109-11118.

Brook RD, Cakmak S, Turner MC, Brook JR, Crouse DL, Peters PA, et al. 2013. Long-term fine particulate matter exposure and mortality from diabetes in Canada. Diabetes Care 36:3313-3320.

Buckland ST, Burnham KP, Augustin NH. 1997. Model selection: An integral part of inference. Biometrics 53:603-618.

Calle E, Terrell D. 1993. Utility of the national death index for ascertainment of mortality among cancer prevention study ii participants. American J of Epidemiol 137:235-241.

Chen H, Burnett RT, Kwong JC, Villeneuve PJ, Goldberg MS, Brook RD, et al. 2013. Risk of incident diabetes in relation to long-term exposure to fine particulate matter in Ontario, Canada. Environ Health Perspect 121:804-810.

Crouse DL, Peters PA, van Donkelaar A, Goldberg MS, Villeneuve PJ, Brion O, et al. 2012. Risk of non accidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: A canadian national-level cohort study. Environmen Health Perspect 120:708-714.

Crouse DL, Peters PA, Hystad P, Brook JR, van Donkelaar A, Martin RV et al. 2015Ambient $PM_{2.5}$, O_3, and NO_2 Exposures and associations with mortality over 16 years of follow-up in the Canadian Census Health and Environment Cohort (CanCHEC). Environ Health Perspect. 123:1180-6.

Evans J, van Donkelaar A, Martin RV, Burnett R, Rainham DG, Birkett NJ, et al. 2013. Estimates of global mortality attributable to particulate air pollution using satellite imagery. Environ Res 120:33-42.

Eze IC, Hemkens LG, Bucher HC, Hoffmann B, Schindler C, Künzli N, Schikowski T, Probst-Hensch NM. 2015. Association between ambient air pollution and diabetes mellitus in Europe and North America: systematic review and meta-analysis. Environ Health Perspect. 123:381-9.

Faes C, Aerts M, Geys H, Molenberghs G. 2007. Model averaging using fractional polynomials to estimate a safe level of exposure. Risk Anal 27:111-123.

Hoek G, Krishnan RM, Beelen R, Peters A, Ostro B, Brunekreef B, Kaufman JD. 2013. Long-term air pollution exposure and cardio- respiratory mortality: a review. Environ Health. 28: 43.

Hsu NC, Jeong MJ, Bettenhausen C, Sayer AM, Hansell R, Seftor CS, et al. 2013. Enhanced deep blue aerosol retrieval algorithm: The second generation. J Geophys Res -Atmos 118: 9296-9315.

Intergovernmental Panel on Climate Change 2014. Climate change 2013: The physical science basis: Working group I contribution to the fifth assessment report of the Intergovernmental Panel on Climate Change: Cambridge University Press.

Jerrett M, Burnett RT, Ma RJ, Pope CA, Krewski D, Newbold KB, et al. 2005. Spatial analysis of air pollution and mortality in Los Angeles. Epidemiology 16:727-736.

Jerrett M, Burnett RT, Pope CA, II, Ito K, Thurston G, Krewski D, et al. 2009. Long-term ozone exposure and mortality. N EnglJ Med 360:1085-1095.

Jerrett M, Burnett RT, Beckerman BS, Turner MC, Krewski D, Thurston G, et al. 2013. Spatial analysis of air pollution and mortality in California. Am J Respir Crit Care Med 188:593-599.

Kloog I, Melly SJ, Ridgway WL, Coull BA, Schwartz J. 2012a. Using new satellite based exposure methods to study the association between pregnancy $PM_{2.5}$ exposure, premature birth and birth weight in Massachusetts. Environ Health 11:40.

Kloog I, Nordio F, Coull BA, Schwartz J. 2012b. Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal pm2.5 exposures in the mid-atlantic states. Environ Sci Technol 46:11913-11921.

Kloog I, Ridgway B, Koutrakis P, Coull BA, Schwartz JD. 2013. Long- and short-term exposure to $PM_{2.5}$ and mortality: Using novel exposure models. Epidemiology 24:555-561.

Krewski D, Burnett RT, Goldberg MS, Hoover K, Siemiatycki J, Jerrett M, et al. 2000. Reanalysis of the Harvard six cities study and the American Cancer Society study of particulate air pollution and mortality. (HEI Special Report). Cambridge, MA:Health Effects Institute.

Krewski D, Jerrett M, Burnett RT, Ma R, Hughes E, Shi Y, et al. 2009. Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality. Research report (Health Effects Institute):5-114; discussion 115-136.

Lee S-J, Serre ML, van Donkelaar A, Martin RV, Burnett RT, Jerrett M. 2012. Comparison of geostatistical interpolation and remote sensing techniques for estimating long-term exposure to ambient $PM_{2.5}$ concentrations across the continental United States. Environ Health Perspect 120:1727-1732.

Lee M, Kloog I, Chudnovsky A, Lyapustin A, Wang Y, Melly S et al. 2015. Spatiotemporal prediction of fine particulate matter using high-resolution satellite images in the Southeastern US 2003-2011. J Expo Sci Environ Epidemiol Jun 17. doi: 10.1038/jes.2015.41. [Epub ahead of print] PubMed PMID: 26082149.

Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: A systematic analysis for the global burden of disease study 2010. Lancet 380:2224-2260.

Lyapustin A, Wang Y, Laszlo I, Kahn R, Korkin S, Remer L, et al. 2011. Multiangle implementation of atmospheric correction (maiac): 2. Aerosol algorithm. J Geophys Res -Atmos 116.

Ma Z, Hu X, Huang L, Bi J, Liu Y. 2014. Estimating ground-level $PM_{2.5}$ in China using satellite remote sensing. Environ Sci Technol 48:7436-7444.

Madrigano J, Kloog I, Goldberg R, Coull BA, Mittleman MA, Schwartz J. 2013. Long-term exposure to $PM_{2.5}$ and incidence of acute myocardial infarction. Environment Health Perspect 121:192-196.

McMillan NJ, Holland DM, Morara M, Feng J. 2010. Combining numerical model output and particulate data using bayesian space–time modeling. Environmetrics 21:48-65.

Molitor J, Jerrett M, Chang CC, Molitor NT, Gauderman J, Berhane K, et al. 2007. Assessing uncertainty in spatial exposure models for air pollution health effects assessment. Environ Health Perspect. 115:1147-53.

Paciorek CJ, Liu Y. 2009. Limitations of remotely sensed aerosol as a spatial proxy for fine particulate matter. Environ Health Perspect 117:904-909.

Pope CA, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, et al. 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA 287:1132-1141.

Pope CA III, Turner MC, Burnett RT, Jerrett M, Gapstur SM, Diver WR, Krewski D, Brook RD. 2015. Relationships between fine particulate air pollution, cardiometabolic disorders, and cardiovascular mortality. Circ Res. 116:108-15.

Rao X, Patel P, Puett R, Rajagopalan S. 2015. Air pollution as a risk factor for type 2 diabetes. Toxicol Sci 143: 231-41.

Snider, G, C. L. Weagle, R. V. Martin, A. van Donkelaar, K. Conrad, D. Cunningham, M. et al. 2015. SPARTAN: A global network to evaluate and enhance satellite-based estimates of ground-level particulate matter for global health applications. Atmos Meas Tech 8: 505-521.

Thurston GD, Ahn J, Cromar KR, Shao Y, Reynolds HR, Jerrett M, Lim CC, Shanley R, Park Y, Hayes RB. Ambient particulate matter air pollution exposure and mortality in the NIH-AARP Diet and Health Cohort. 2015. Environmen Health Perspect. Sep 15. [Epub ahead of print] DOI:10.1289/ehp.1509676.

Turner MC, Cohen A, Jerrett M, Gapstur SM, Diver WR, Pope CA, et al. 2014. Interactions between cigarette smoking and fine particulate matter in the risk of lung cancer mortality in Cancer Prevention Study II. Am J Epidemiol 180:1145-9.

Turner MC, Jerrett M, Pope Iii CA, Krewski D, Gapstur SM, Diver WR, Beckerman BS, Marshall JD, Su J, Crouse DL, Burnett RT. 2015. Long-term ozone exposure and mortality in a large prospective study. Am J Respir Crit Care Med. First published online 17 Dec 2015 as DOI: 10.1164/rccm.201508-1633OC.

van Donkelaar A, Martin RV, Brauer M, Kahn R, Levy R, Verduzco C, et al. 2010. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: Development and application. Environ Health Perspect 118:847-855.

van Donkelaar A, Martin RV, Brauer M, Boys BL. 2015a. Global fine particulate matter concentrations from satellite for long-term exposure assessment. Environ Health Perspect 123: 135-143.

van Donkelaar, A., R. V. Martin, R. J. D. Spurr and R. T. Burnett. 2015b. High-resolution satellite-derived $PM_{2.5}$ from optimal estimation and geographically weighted regression over North America, Environ Sci and Tech doi 10.1021/acs.est.5b02076.

van Donkelaar, A., R.V Martin, M.Brauer, N. C. Hsu, R. A. Kahn, R. C Levy, A. Lyapustin, A. M. Sayer, and D. M Winker, Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors, ***Environ. Sci. Technol.***, doi: 10.1021/acs.est.5b05833, 2016.

Vienneau D, de Hoogh K, Bechle MJ, Beelen R, van Donkelaar A, Martin RV, et al. 2013. Western european land use regression incorporating satellite- and ground-based measurements of NO_2 and PM_{10}. Environ Sci Technol 47:13555-13564.

Winker DM, Vaughan MA, Omar A, Hu Y, Powell KA, Liu Z, et al. 2009. Overview of the Calipso mission and Caliop data processing algorithms. J Atmos and Oceanic Technol 26:2310-2323.

TABLES

Table 1. Model descriptions including spatial and temporal dimensions, auxiliary data, and cross-validation summary

Model Name	Model Type	Spatial Scale	Temporal Scale	Ground Data Used	Other Auxiliary Data	Cross-validation Methods and Results
PM2.5 HBMCMAQ 02-04	Atmospheric Chemistry with Statistical Data Fusion	36 km * 36 km grid	2002-2004	Yes used in the data fusion.	Yes, meteorological data.	Conducted for sub-area of the Northeaster and Midwestern parts of the continent with a 12 km grid; 44 Federal Reference Method sites used for cross-validation results found to track monitoring data temporal patterns well, but with some seasonal bias. Results outperformed exponential kriging model for bias and were slightly worse for Mean Square Error (MSE) (McMillian et al. 2010).
PM2.5 BME 02-04	Bayesian Maximum Entropy Space-time kriging	Predicted at centroids of 0.1° × 0.1° ~9.8 * 9.8 km grid used for estimating the PM2.5 RS 01-06 (see below)	2002-2004	Y based on 1318 monitors with monthly averages	No	Extensive cross-validation based on 146 leave-out sites with MSE generally less 5 for distances less than 98.7 km from the cross-validation site this model predicted ground level concentrations more accurately than the $PM_{2.5}$ RS 01-06 , while at greater distances than this, the $PM_{2.5}$ RS 01-06 outperformed the kriging model in terms of MSE (see below) (Lee et al. 2012).
PM2.5 BMELUR 02-04	Land use regression model with Bayesian Maximum Entropy kriging of residuals from the LUR model	30 * 30 m estimate centered on the target receptor location	2002-2004	Y based on 1318 monitors with monthly averages	Yes, traffic density within 1 km of the monitor and open space within 400 m of the monitor	Model variables selected with a machine-learning algorithm that used v-fold cross-validation in the model selection. About 10% of data held out for cross-validation (i.e., 146 ground sites). Cross-validation $r^2 \sim 0.8$ with little evidence of bias or heterskewdasticity (Beckerman et al. 2012).
PM2.5BMELURRS 02-04	Land use regression model with Bayesian	30*30 m estimate centered on the target	2002-2004	Y based on 1318 monitors with	Yes, remote sensing estimate at ~9.8 km and	Model underwent the same cross-validation as the $PM_{2.5}$ BMELUR 02-04

	Maximum Entropy kriging of residuals from	receptor location		monthly averages	developed land use at XX m	above. Cross-validation r^2 ~ 0.8 with no apparent sign of bias or heterskewdasticity (Beckerman et al. 2012)
$PM_{2.5}$ GWR RS 02-04	Based on AOD from the MODIS satellite instrument scaled to prior years using AOD from the MISR and SeaWiFS satellite instruments and adjusted with geographically weighted regression	0.01° × 0.01°~ 1 × 1 km grid	2002-2004	Yes, used in the geographically weighted regression	Atmospheric Chemical Transport Model (GEOS-Chem); scaling to years prior to 2004 follows van Donkelaar et al., 2015a.	r^2=0.79 (cross-validated); Uncertainty=N(-0.38,1.49) μ g/m^3; RMSD = 1.5 μ g/m^3; based on comparison with 2004-2008 $PM_{2.5}$ observed at 1,440 ground monitors (van Donkelaar et al. 2015b)
$PM_{2.5}$ No GWR RS 02-04	Based on AOD from the MODIS satellite instrument scaled to prior years using AOD from the MISR and SeaWiFS satellite instruments	0.01° × 0.01°~ 1 × 1 km grid	2002-2004	No	Atmospheric Chemical Transport Model (GEOS-Chem); scaling to years prior to 2004 follows van Donkelaar et al., 2015a.	r^2=0.62; Uncertainty=N(-0.87,2.42) μ g/m^3; RMSD = 2.6.5 μ g/m^3; based on comparison with 2004-2008 $PM_{2.5}$ observed at 1,440 ground monitors (van Donkelaar et al. 2015b)
PM2.5 RS 01-06	Based on aerosol optical depth from the MODIS and MISR satellite instruments	0.1° × 0.1°~ 9.8 * 9.8 km grid	2001-2006	No	Atmospheric Chemical Transport Model (GEOS-Chem)	r^2=0.49 (non-coincident); r^2=0.59; 1-σ error = 1 μg/m^3 + 15%; y=1.07x-1.75 (coincident) based on comparison with 2001-2006 $PM_{2.5}$ observed at 1,057 ground monitors (van Donkelaar et al., 2010; 2015)

Table 2. Descriptive statistics for the exposure models after assignment to ACS CPS II participants.

Air Pollution	N	Mean (SD)	Minimum	10th percentile	1st quartile	2nd quartile	3rd quartile	90th percentile	Maximum	IQR	Range
PM2.5 HBMCMAQ 02-04	668,629	12.1 (2.6)	2.8	8.7	10.4	12.1	14.0	15.2	21.4	3.6	18.6
PM2.5 BME 02-04	668,629	12.1 (2.6)	3.4	8.4	10.3	12.2	13.9	15.0	21.6	3.7	18.2
PM2.5 BMELUR 02-04	668,629	12.0 (2.7)	1.5	8.6	10.1	12.0	13.7	15.1	26.6	3.6	25.1
PM2.5 BMELUR CT 02-04	668,629	11.7 (2.8)	1.0	8.2	9.8	11.7	13.5	14.9	26.2	3.7	25.2
PM2.5 BMELURRS 02-04	668,629	12.0 (2.8)	3.2	8.4	10.0	11.9	13.8	15.2	24.4	3.7	21.2
PM2.5 BMELURRS CT 02-04	668,629	11.8 (2.8)	2.8	8.1	9.8	11.8	13.6	15.1	24.4	3.8	21.6
PM2.5 RS GWR CT 02-04	668,629	12.2 (3.2)	1.3	7.9	9.9	12.7	14.6	16.0	25.4	4.7	24.1
PM2.5 RS no GWR CT 02-04	668,629	11.4 (3.6)	0.7	6.1	8.6	12.1	14.2	15.7	22.5	5.6	21.8
PM2.5 BMELUR 01-06	668,629	12.1 (2.6)	1.4	8.7	10.2	12.1	13.9	15.2	25.8	3.8	24.4
PM2.5 RS 01-06	668,629	11.9 (3.8)	1.9	7.0	9.0	11.8	14.7	16.9	24.6	5.7	22.6

Table 3. Correlations among the estimates of $PM_{2.5}$ after assignment to ACS CPS II participants.

Air Pollution	PM2.5 HBMCMAQ 02-04	PM2.5 BME 02-04	PM2.5 BMELUR 02-04	PM2.5 BMELUR CT 02-04	PM2.5 BMELUR RS 02-04	PM2.5 BMELURRS CT 02-04	PM2.5 RS GWR CT 02-04	PM2.5 RS no GWR CT 02-04	PM2.5 BMELUR 01-06	PM2.5 RS 01-06
PM2.5 HBMCMAQ 02-04	1.00	0.88	0.84	0.82	0.85	0.85	0.71	0.59	0.84	0.63
PM2.5 BME 02-04		1.00	0.92	0.90	0.93	0.93	0.73	0.62	0.92	0.64
PM2.5 BMELUR 02-04			1.00	0.94	0.94	0.93	0.71	0.55	0.99	0.60
PM2.5 BMELUR CT 02-04				1.00	0.91	0.93	0.72	0.54	0.94	0.60
PM2.5 BMELURRS 02-04					1.00	0.99	0.72	0.58	0.93	0.66
PM2.5 BMELURRS CT 02-04						1.00	0.73	0.58	0.92	0.67
PM2.5 RS GWR CT 02-04							1.00	0.78	0.74	0.72
PM2.5 RS no GWR CT 02-04								1.00	0.59	0.84
PM2.5 BMELUR 01-06									1.00	0.62
PM2.5 RS 01-06										1.00

Table 4. Results of the Cox proportional hazard modeling with adjustment for individual or individual plus year 1990 ecologic covariates. Hazard ratios expressed over a 10 μg/m^3 increment.

	Diseases of the circulatory system				Ischemic heart disease			
	N=100,102				N=45,624			
Air Pollution	Fully-adjusted HR (95% CI)	AIC (1,587,000s)	Fully-adjusted HR (95% CI) + 1990 Ecological Confounders	AIC (1,587,000s)	Fully-adjusted HR (95% CI)	AIC (726,000s)	Fully-adjusted HR (95% CI) + 1990 Ecological Confounders	AIC (726,000s)
PM2.5 HBMCMAQ 02-04	1.09 (1.07-1.12)	434	1.09 (1.06-1.12)	094	1.15 (1.11-1.19)	688	1.11 (1.07-1.16)	315
PM2.5 BME 02-04	1.13 (1.10-1.15)	388	1.12 (1.09-1.15)	065	1.19 (1.15-1.23)	650	1.15 (1.10-1.19)	296
PM2.5 BMELUR 02-04	1.15 (1.13-1.18)	340	1.14 (1.11-1.17)	033	1.20 (1.16-1.24)	636	1.15 (1.11-1.19)	290
PM2.5 BMELUR CT 02-04	1.13 (1.11-1.16)	364	1.12 (1.09-1.15)	051	1.19 (1.15-1.23)	643	1.14 (1.10-1.18)	292
PM2.5 BMELURRS 02-04	1.12 (1.09-1.14)	388	1.11 (1.08-1.14)	066	1.18 (1.14-1.22)	652	1.13 (1.09-1.17)	297
PM2.5 BMELURRS CT 02-04	1.11 (1.09-1.13)	396	1.11 (1.08-1.13)	068	1.17 (1.13-1.20)	660	1.12 (1.08-1.16)	301
PM2.5 RS GWR CT 02-04	1.09 (1.07-1.11)	411	1.08 (1.06-1.11)	088	1.10 (1.06-1.13)	711	1.08 (1.05-1.12)	321
PM2.5 RS no GWR CT 02-04	1.04 (1.03-1.06)	462	1.02 (1.00-1.04)	131	1.09 (1.06-1.12)	707	1.06 (1.02-1.09)	331
PM2.5 BMELUR 01-06	1.16 (1.13-1.19)	336	1.14 (1.11-1.17)	036	1.20 (1.16-1.25)	639	1.15 (1.11-1.19)	293
PM2.5 RS 01-06	1.05 (1.04-1.07)	447	1.05 (1.03-1.07)	115	1.12 (1.10-1.15)	658	1.10 (1.07-1.14)	298
Ensemble Estimate	1.16 (1.13-1.19)	NA	1.14 (1.11-1.17)	NA	1.20 (1.16-1.24)	NA	1.15 (1.11-1.19)	NA

Note 1: There are 43 variables in the model including $PM_{2.5}$ for individual only and 55 in fully adjusted.

Table 5. Cox proportional hazard model results restricted to those participants residing in a metropolitan area with a central monitor measurement of pollution. Hazard ratios expressed over a 10 $\mu g/m^3$ increment.

	Diseases of the circulatory system		Ischemic heart disease	
Air Pollution	Fully-adjusted HR (95% CI) + 1990 Ecological Confounders	AIC (801,000s)	Fully-adjusted HR (95% CI) + 1990 Ecological Confounders	AIC (373,000s)
Central Monitor 99-00	1.08 (1.05-1.12)	533	1.11 (1.06-1.16)	292
PM2.5 HBMCMAQ 02-04	1.09 (1.05-1.14)	536	1.09 (1.03-1.16)	301
PM2.5 BME 02-04	1.11 (1.07-1.15)	526	1.12 (1.06-1.18)	294
PM2.5 BMELUR 02-04	1.12 (1.09-1.16)	510	1.12 (1.06-1.18)	291
PM2.5 BMELUR CT 02-04	1.12 (1.08-1.16)	512	1.12 (1.07-1.18)	289
PM2.5 BMELURRS 02-04	1.10 (1.06-1.13)	527	1.11 (1.06-1.16)	293
PM2.5 BMELURRS CT 02-04	1.09 (1.06-1.13)	527	1.10 (1.05-1.15)	295
PM2.5 RS GWR CT 02-04	1.08 (1.04-1.12)	539	1.09 (1.04-1.15)	300
PM2.5 RS no GWR CT 02-04	1.00 (0.97-1.03)	555	1.03 (0.98-1.07)	309
PM2.5 BMELUR 01-06	1.13 (1.09-1.07)	511	1.12 (1.07-1.18)	291
PM2.5 RS 01-06	1.03 (1.00-1.06)	552	1.10 (1.06-1.15)	288

Total n = 379618 with 54689 deaths from circulatory disease and 25393 from ischemic heart disease.
Note 1: There are 43 variables in the model including $PM_{2.5}$ for individual only and 55 in fully adjusted.