Journal of Exposure Science and Environmental Epidemiology (2011) 21, 10–19

www.nature.com/jes

The use of alternative pollutant metrics in time-series studies of ambient air pollution and respiratory emergency department visits

LYNDSEY A. DARROW[a], MITCHEL KLEIN[a], JEREMY A. SARNAT[a], JAMES A. MULHOLLAND[b], MATTHEW J. STRICKLAND[a], STEFANIE E. SARNAT[a], ARMISTEAD G. RUSSELL[b] AND PAIGE E. TOLBERT[a]

[a]*Emory University, Atlanta, Georgia, USA*
[b]*Georgia Institute of Technology, Atlanta, Georgia, USA*

Various temporal metrics of daily pollution levels have been used to examine the relationships between air pollutants and acute health outcomes. However, daily metrics of the same pollutant have rarely been systematically compared within a study. In this analysis, we describe the variability of effect estimates attributable to the use of different temporal metrics of daily pollution levels. We obtained hourly measurements of ambient particulate matter ($PM_{2.5}$), carbon monoxide (CO), nitrogen dioxide (NO_2), and ozone (O_3) from air monitoring networks in 20-county Atlanta for the time period 1993–2004. For each pollutant, we created (1) a daily 1-h maximum; (2) a 24-h average; (3) a commute average; (4) a daytime average; (5) a nighttime average; and (6) a daily 8-h maximum (only for O_3). Using Poisson generalized linear models, we examined associations between daily counts of respiratory emergency department visits and the previous day's pollutant metrics. Variability was greatest across O_3 metrics, with the 8-h maximum, 1-h maximum, and daytime metrics yielding strong positive associations and the nighttime O_3 metric yielding a negative association (likely reflecting confounding by air pollutants oxidized by O_3). With the exception of daytime metric, all of the CO and NO_2 metrics were positively associated with respiratory emergency department visits. Differences in observed associations with respiratory emergency room visits among temporal metrics of the same pollutant were influenced by the diurnal patterns of the pollutant, spatial representativeness of the metrics, and correlation between each metric and copollutant concentrations. Overall, the use of metrics based on the US National Ambient Air Quality Standards (for example, the use of a daily 8-h maximum O_3 as opposed to a 24-h average metric) was supported by this analysis. Comparative analysis of temporal metrics also provided insight into underlying relationships between specific air pollutants and respiratory health.
Journal of Exposure Science and Environmental Epidemiology (2011) **21,** 10–19; doi:10.1038/jes.2009.49; published online 16 September 2009

Keywords: *air pollution, respiratory disease, emergency department visits, exposure assessment, criteria pollutants, time-series.*

Introduction

Studies of the acute health effects of ambient air pollution have used various temporal metrics to characterize daily carbon monoxide (CO), nitrogen dioxide (NO_2), ozone (O_3), and fine particulate matter ($PM_{2.5}$) concentrations. Commonly examined metrics include 24-h average concentrations, daily 1-h maximum concentrations, and daily 8-h maximum concentrations. Bell et al. (2005) showed that variation in O_3 between cities and over time, differed by averaging time used to characterize the O_3 concentrations (that is, 24-h average, 1-h maximum, and 8-h maximum). The choice of daily pollutant averaging time could likewise affect risk estimates observed in epidemiological studies. Some pollutant averaging times may show stronger associations with health outcomes, because they reflect a more biologically relevant exposure (for example, peak *vs* average exposure) or because they more strongly correlate with average population exposures compared with other temporal metrics. In addition, certain metrics could act as better surrogates for other, unmeasured pollutants responsible for the adverse health effects, such as certain metrics of CO and NO_2 acting as surrogates of particles from traffic sources (Sarnat et al., 2001).

Temporal metrics that reflect peak pollution levels (for example, 1-h maximum) may be the most biologically relevant if the health effect is triggered by a high, short-term dose rather than a steady dose throughout the day. Peak concentrations, however, are frequently associated with episodic, local emission events, resulting in spatially heterogeneous concentrations across an urban area and, thus, prone to measurement error when using fixed-site concentrations as the estimate of exposure. As a result, a 24-h average concentration metric may often be more representative of average population exposures.

It is also possible that the most appropriate temporal metric for an epidemiological analysis is determined by exposure factors related to population time–activity patterns; some metrics may better capture average population exposures, because they include hours when the population

1. Address all correspondence to: Dr. Lyndsey A. Darrow, Department of Environmental Health, Rollins School of Public Health of Emory University, 1518 Clifton Road, NE–Room 240, Atlanta, GA 30322, USA. Tel.: +404 727 4595. Fax: +404 727 8744.
E-mail: ldarrow@sph.emory.edu
Received 12 April 2009; accepted 30 July 2009; published online 16 September 2009

is most likely to be exposed to ambient air. A relevant exposure time window for assessing health effects of traffic pollution, for example, may be during heavy commuting hours, when pollutant concentrations are highest and people are more directly exposed to ambient air. Alternatively, if centrally located monitoring stations, which are often used in epidemiological studies to characterize exposure, only reflect downtown concentrations, daytime hours might correlate best with personal exposures given the influx of people into the city center during the day and exodus at night. Conversely, nighttime exposure metrics incorporate hours when people are likely to be in their homes and less likely to be outdoors and exposed directly to ambient air, thus increasing exposure measurement error. As exposure measurement error can lead to attenuated effect estimates, if some pollutant metrics approximate population exposures better than other metrics, we would expect to see variation in epidemiological results according to choice of metric. Furthermore, if associations between nighttime concentrations and health outcomes are dramatically different from associations between daytime concentrations and health outcomes, this may indicate that a 24-h average metric is inappropriate, as it inherently combines both day and night concentrations into one metric. Lastly, if the use of different temporal metrics of air pollution leads to different results, metric choice could potentially explain differences in observed epidemiological associations across studies.

In our previous analyses of ambient air pollution levels and respiratory emergency department visits in Atlanta (Metzger et al., 2004; Peel et al., 2005; Sarnat et al., 2008; Tolbert et al., 2007; Sarnat et al., 2009), we presented results using *a priori* exposure metrics chosen for each pollutant of interest, based on the National Ambient Air Quality Standards and previous studies of air pollution and acute health effects (Sunyer et al., 1997; Zmirou et al., 1998; Ostro et al., 2001). These *a priori* exposure metrics included the daily 1-h maximum concentration for CO and NO_2, the daily 8-h maximum concentration for O_3, and a 24-h average for $PM_{2.5}$. In the present study, we describe the variability of epidemiological results attributable to the use of different daily metrics of the same pollutant, comparing the results using our *a priori* pollutant metrics to those obtained using alternative temporal metrics of pollutant concentrations. The degree of sensitivity to the choice of metric is relevant to future investigations of air pollution and respiratory health, can offer clues to biological mechanisms, and ultimately can be used to inform regulatory policy.

Methods

Air Quality Data

We obtained hourly ambient concentrations of CO, NO_2, O_3, and $PM_{2.5}$ from the US Environmental Protection Agency's Air Quality System as well as the Aerosol Research Inhalation Epidemiology Study monitor located near downtown Atlanta (Van Loy et al., 2000; Hansen et al., 2006). Data from all monitoring stations in the study area that provided hourly measurements were used to assess the spatial heterogeneity of each metric (described below). However, for the epidemiological models, we used measurements from a single, centrally located monitor for each pollutant. We obtained daily meteorological data from the National Climatic Data Center at Hartsfield–Atlanta International Airport. We chose to examine CO, NO_2, and O_3 because these pollutants were associated with respiratory emergency room visits in our previous analyses (Peel et al., 2005). We also assessed $PM_{2.5}$ because our previous analyses were suggestive of an effect in spite of limited sample size (Peel et al., 2005). Furthermore, a motivation of the study was to explore whether alternative temporal metrics would yield stronger associations than our *a priori* daily metrics, which are metrics commonly used in air pollution studies. For each pollutant, we created the following temporal metrics of daily pollutant concentrations: a daily1-h maximum, a 24-h average, an average of commute hours ("commute," 0700–1000 hours and 1600–1900 hours), a daytime average ("daytime," 0800–1900 hours), a nighttime average ("nighttime," 2400–0600 hours), and, for ozone only, a daily 8-h maximum. Within a pollutant, the study days included in the analysis were the same across metrics. The analytic time period differed by pollutant depending on the period of monitoring available at the central monitoring station: CO was examined from 1 January 1993 through to 30 June 2003; NO_2 was examined between 1 March 1994 and 31 December 2004; O_3 was examined from March through to October of every year between 1993 and 2004; $PM_{2.5}$ was examined from 1 August 1998 through to 31 December 2004.

Emergency Department Data

Individual-level data from computerized billing records were obtained from 41 of 42 acute-care hospitals in the 20-county Atlanta area (50-mile radius). We examined daily counts of selected respiratory-related emergency department visits for patients living within any one of the 225 ZIP codes located wholly or partially in the 20-county Atlanta study area. Emergency department visits with a primary International Classification of Diseases 9th Revision (ICD-9) diagnostic code for asthma and wheeze (493, 786.09), chronic obstructive pulmonary disease (491, 492, 496), upper respiratory infection (460–466, 477), and pneumonia (480–486) were classified as respiratory emergency department visits. We excluded repeat visits by patients visiting the same hospital within a single day. There were 1,068,525 respiratory emergency department visits between 1993 and 2004, with an average of 244 visits per day.

Statistical Analysis
We modeled the association between air pollution and respiratory emergency department visits using a case-crossover framework, a special case of the time-series approach (Lu and Zeger, 2007). Using this time-stratified approach, referent days were chosen within the same calendar month and within the same maximum temperature category as the day of the respiratory emergency department visit (Schwartz, 2004, 2005; Zanobetti and Schwartz, 2006). For example, if the visit occurred in March 2000 on a day with a maximum temperature of 72°F (the case day), we selected all other days in March 2000 with a maximum temperature between 70°F and 75°F as the control days. Maximum temperature categories were in five degree increments and three degree increments at the extremes: <35°F or >89°F. Counts were then pooled across individuals within a hospital to create a time series of counts for each hospital. We chose to match on temperature rather than day-of-week because temperature effects are non-linear and can be challenging to adequately control in regression models compared with day-of week, which can be controlled using indicator variables.

We analyzed the data using Poisson generalized linear models, scaling the variance estimates to account for overdispersion. The model included indicator variables for day-of-week and holidays, cubic terms for 2-day moving average (lag 1–2) minimum temperature (same-day temperature was accounted for by matching) and for 3-day moving average (lag 0–2) of dew point temperature (cubic terms). We also repeated the analysis using the time-series models from our previously published work (Peel et al., 2005) to evaluate whether the observed patterns were sensitive to modeling approach. Briefly, our previous time-series models included cubic splines with monthly knots to control for temporal trends, seasonal indicator variables, and cubic splines to control for temperature and dew point temperature. All analyses were performed using SAS, version 9.2, statistical software (SAS Institute, Cary, North Carolina, USA).

Metrics Comparison
We calculated partial correlations (that is, correlations after controlling for the covariates included in the time-series models) between all of the metrics, both within and across pollutants. We compared the spatial heterogeneity of the metrics for each pollutant to assess whether some metrics might better reflect population-wide exposures (Ito et al., 2001); metrics that are more spatially representative (that is, more correlated across space) might better reflect personal exposures in the study population, thus reducing bias due to measurement error relative to other metrics. Thus, comparing the spatial heterogeneity of the metrics may shed light on any observed differences in strength of association with respiratory emergency room visits. To compare the spatial heterogeneity of the different averaging times for a given pollutant, we created the metrics at every air quality monitoring station in the study area that measured hourly concentrations. As these additional monitoring stations were located at various distances from the central monitor, we were able to assess the degree of correlation at various distances for each of the metrics.

We examined associations between daily respiratory emergency department visits and metrics of CO, NO_2, O_3, and $PM_{2.5}$ concentrations on the previous day, lag 1. In the primary analyses, we chose to focus on previous day (lag 1) pollution concentrations, as this lag was consistently associated with the outcome in previous analyses (Peel et al., 2005). Owing to the uncertainty of the relevant lag period of exposure for the pollutants of interest, in sensitivity analyses, we also examined alternative lags of each metric (lag days 0, 2, 3). Lag 0 was defined as the period from midnight to midnight on the day of the visit; lag 1 was defined as the period from midnight to midnight on the day preceding the visit, and so on. To compare the magnitude of effect across different metrics of the same pollutant, we calculated risk ratios for both an interquartile range (IQR) increase in concentration, which differed across metrics, and for a standard unit increase, which was the same for all metrics of a given pollutant. The risk ratios for an IQR increase allowed for a comparison of effects for a similar degree of increase relative to each metric's distribution of concentrations, whereas the risk ratios for a standard unit increase allowed us to compare the magnitude of effect for an absolute increase (for example, 10 p.p.b.) in pollutant concentration. χ^2-values and corresponding P-values, which are not unit dependent, were calculated to compare the strength of statistical association for each pollutant metric. χ^2- and P-values, which are highly influenced by sample size, could be compared because the number of days included in the analysis was the same across metrics for a given pollutant.

Results

Descriptive statistics for each of the pollutant metrics examined are presented in Table 1. For a given pollutant, many of the metrics were highly correlated (Table 2). As expected, correlations were generally higher between overlapping temporal metrics; for example, the nighttime and daytime metrics were less correlated with each other than with the 24-h average, which encompassed both nighttime and daytime hours. For CO, correlations among the metrics ranged from 0.48 to 0.91, with the weakest correlation observed between the daytime and nighttime values. Correlations among NO_2 metrics ranged from 0.44 to 0.90, with relatively weak correlations between the nighttime and daytime metrics and between the 1-h maximum and daytime metrics ($r = 0.45$ and 0.44, respectively). The O_3 metrics were more strongly correlated with one another ($r = 0.68$–0.95) with the exception of the nighttime metric, which was

Table 1. Descriptive statistics for air pollution data.

Pollutant	N^a	Time period	Metric	Mean	SD	Percentiles			
						25th	50th	75th	Max
CO (p.p.m.)	3486	1/1/1993–6/30/2003	1-h maximum	1.6	1.1	0.8	1.3	2.2	7.7
			24-h average	0.7	0.4	0.5	0.6	0.9	3.7
			Commute (0700–1000 hours, 1600–1900 hours)	0.7	0.4	0.4	0.6	0.9	4.0
			Day-time (0800–1900 hours)	0.5	0.3	0.4	0.5	0.7	2.6
			Night-time (2400–0600 hours)	0.8	0.7	0.3	0.5	1.0	5.2
NO_2 (p.p.b.)	3635	3/1/1994–12/31/2004	1-h maximum	43	18	30	41	53	181
			24-h average	22	10	15	21	28	74
			Commute (0700–1000 hours, 1600–1900 hours)	21	11	13	20	27	97
			Day-time (0800–1900 hours)	17	9	10	16	22	82
			Night-time (2400–0600 hours)	25	16	13	22	35	97
O_3 (p.p.b.)	2883	March–October, 1993–2004	8-h maximum	53	22	38	51	67	148
			1-h maximum	62	25	45	59	76	180
			24-h average	30	12	21	29	37	81
			Commute (0700–1000 hours, 1600–1900 hours)	35	16	24	35	45	106
			Day-time (0800–1900 hours)	45	20	32	44	58	123
			Night-time (2400–0600 hours)	14	12	3	11	22	64
$PM_{2.5}$ ($\mu g\,m^{-3}$)	1660	8/1/1998–12/31/2004	1-h maximum	29	16	18	26	36	188
			24-h average	16	9	10	14	21	72
			Commute (0700–1000 hours, 1600–1900 hours)	17	9	10	15	21	76
			Day-time (0800–1900 hours)	15	8	8	13	19	71
			Night-time (2400–0600 hours)	17	11	5	9	14	88

[a]Number of days used in analysis.
$PM_{2.5}$ was measured continuously using a tapered element oscillating microbalance (TEOM) operated at 30°C to minimize volatilization.
Gaseous pollutants were measured using standard approaches (NO_2 and O_3 by chemiluminescence and CO by infrared analyzer).

uncorrelated with the other O_3 metrics except for the 24-h average metric ($r = 0.46$). Correlations among $PM_{2.5}$ metrics ranged from 0.60 to 0.94. Similar to CO and NO_2, the weakest correlation among the $PM_{2.5}$ metrics was between daytime and nighttime.

Diurnal patterns for the traffic-related pollutants (CO, NO_2, and $PM_{2.5}$) were bimodal, with peaks during the morning and evening rush hours (Figure 1). Hourly maxima for CO and NO_2 typically occurred at night between 2100 hours and 2300 hours. Concentrations of these pollutants remain elevated during much of the overnight period because of meteorology. Ozone also exhibited a typical diurnal trend, with peaks occurring in the mid- to late afternoon and minima occurring during the night.

Spatial Correlations of the Metrics

We examined the spatial correlation of various metrics to assess whether differences in spatial correlation between the metrics could explain differences in the observed associations. The spatial correlations between CO, NO_2, O_3, and $PM_{2.5}$ metrics at the various monitoring station distances are shown in Figure 2. Nighttime O_3 was the most spatially heterogeneous of the O_3 metrics; all of the other O_3 metrics showed strong spatial correlations even for long distances between monitors. Spatial correlations for NO_2 were fairly similar across metrics for distances <20 km. However, in comparisons of monitors >38 km apart, the NO_2 daily 1-h maximum was considerably more spatially heterogeneous than the daytime metric. Generally, metrics that included hours when NO_2 concentrations were highest exhibited greater spatial heterogeneity (note that the monitoring station located at 15 km is impacted by a nearby freeway). Similarly, the $PM_{2.5}$ metrics showed strong spatial correlations ($r>0.7$ for all distances) with the exception of the 1-h maximum, which was more spatially heterogeneous ($0.5<r<0.6$ between distances of 10 and 60 km). The 24-h average $PM_{2.5}$ was the most spatially homogeneous of the $PM_{2.5}$ metrics examined.

Comparing Statistical Significance of Association

Risk ratios, 95% confidence intervals, χ^2-values and corresponding P-values from the regression models are shown in Table 3. On the basis of the χ^2-values, the nighttime metrics for both CO and NO_2 were the most strongly associated with respiratory emergency department visits. The daytime metric for CO and NO_2, which not only corresponds to hours of lower concentrations but also periods of time when people are more likely to be exposed, showed the weakest associations for these pollutants. With the exception of the daytime metrics, associations for the CO and NO_2 metrics were all statistically significant.

The daily 1 and 8-h maximums yielded the strongest associations for O_3. The daytime metric, which captured the hours of peak O_3 concentrations, was also strongly

Table 2. Partial[a] spearman correlation coefficients for all air pollutant metrics.

Metric	CO					NO_2					O_3						$PM_{2.5}$			
	1-h	24-h	Com	Day	Night	1-h	24-h	Com	Day	Night	8-h	1-h	24-h	Com	Day	Night	1-h	24-h	Com	Day
CO																				
1-h	1																			
24-h	0.87	1																		
Com	0.65	0.85	1																	
Day	0.53	0.76	0.91	1																
Night	0.53	0.71	0.59	0.48	1															
NO_2																				
1-h	0.61	0.55	0.38	0.28	0.38	1														
24-h	0.62	0.66	0.55	0.44	0.51	0.79	1													
Com	0.47	0.56	0.57	0.49	0.41	0.55	0.84	1												
Day	0.41	0.50	0.54	0.53	0.34	0.44	0.76	0.90	1											
Night	0.47	0.53	0.44	0.31	0.66	0.59	0.78	0.56	0.45	1										
O_3																				
8-h	0.15	0.11	0.02	−0.06	0.14	0.34	0.24	0.12	0.02	0.24	1									
1-h	0.21	0.19	0.10	0.03	0.19	0.40	0.33	0.21	0.13	0.30	0.93	1								
24-h	−0.17	−0.22	−0.22	−0.22	−0.15	0.02	−0.15	−0.17	−0.22	−0.11	0.78	0.68	1							
Com	0.01	−0.07	−0.20	−0.23	−0.02	0.20	0.01	−0.16	−0.21	0.05	0.83	0.74	0.88	1						
Day	0.12	0.06	−0.06	−0.14	0.11	0.31	0.18	0.04	−0.07	0.22	0.95	0.89	0.84	0.91	1					
Night	−0.43	−0.50	−0.38	−0.24	−0.63	−0.35	−0.52	−0.37	−0.31	−0.66	0.04	−0.04	0.46	0.22	0.07	1				
$PM_{2.5}$																				
1-h	0.46	0.48	0.37	0.31	0.39	0.50	0.53	0.46	0.43	0.42	0.39	0.43	0.15	0.22	0.32	−0.23	1			
24-h	0.36	0.45	0.37	0.34	0.45	0.42	0.52	0.47	0.45	0.45	0.46	0.49	0.25	0.31	0.41	−0.19	0.82	1		
Com	0.31	0.40	0.38	0.34	0.40	0.36	0.47	0.46	0.45	0.42	0.42	0.46	0.21	0.26	0.37	−0.19	0.75	0.94	1	
Day	0.21	0.31	0.31	0.33	0.33	0.27	0.37	0.38	0.41	0.33	0.40	0.44	0.24	0.26	0.36	−0.10	0.68	0.91	0.93	1
Night	0.28	0.37	0.31	0.26	0.55	0.33	0.45	0.41	0.37	0.52	0.31	0.32	0.15	0.18	0.28	−0.22	0.62	0.79	0.69	0.60

1-h, 1 h maximum; 24-h, 24 h average; com, commute hours: 0700–1000 hours and 1600–1900 hours; day, workday hours: 0800–1900 hours; night, night hours 2400–0600 hours.
[a]Controlling for covariates included in the Poisson regression model (month–year–maximum temperature strata, lag 1–2 moving average minimum temperature, lag 0–1–2 moving average of dew point temperature, day of week, holidays).

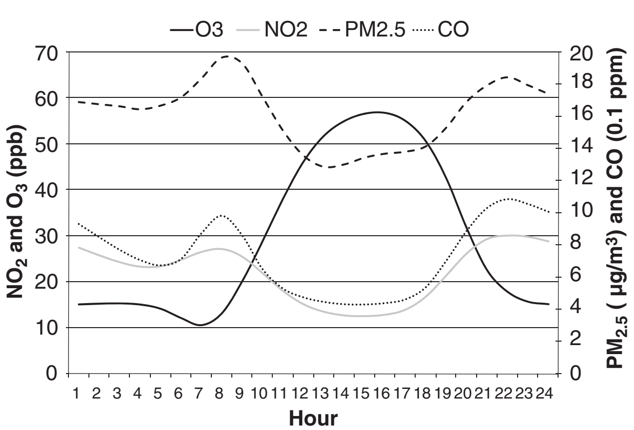

Figure 1. Diurnal pattern* for selected pollutants. *Average of hourly values over study period, hour 1 refers to the hour between 2400 hours and 0100 hours.

associated with respiratory emergency department visits. The commute and the 24-h average metrics for O_3, however, were only weakly associated with respiratory emergency department visits ($P>0.05$), and the nighttime metric of O_3 was inversely associated with respiratory emergency department visits. Ozone was the only pollutant for which the choice of metric affected the direction of association. Given the negative correlations between the nighttime O_3 metric and the CO and NO_2 metrics (Table 2), we suspected that this negative association might be confounded by the positive associations observed with the various metrics of CO and NO_2. In multi-pollutant models, when any of the CO and NO_2 metrics were included in the model as covariates (with the exception of daytime CO), nighttime O_3 was not negatively (or positively) associated with respiratory emergency department visits. There were no observed associations between any of the $PM_{2.5}$ metrics and respiratory emergency department visits. However, the sample size was more limited for $PM_{2.5}$, and all point estimates were above the null; χ^2- values and risk ratios were comparable across metrics.

Table 4 displays the χ^2-values and risk ratios (per interquartile range) for models using alternative lags of each metric (lag days 0, 2, and 3), in addition to the lag 1 day

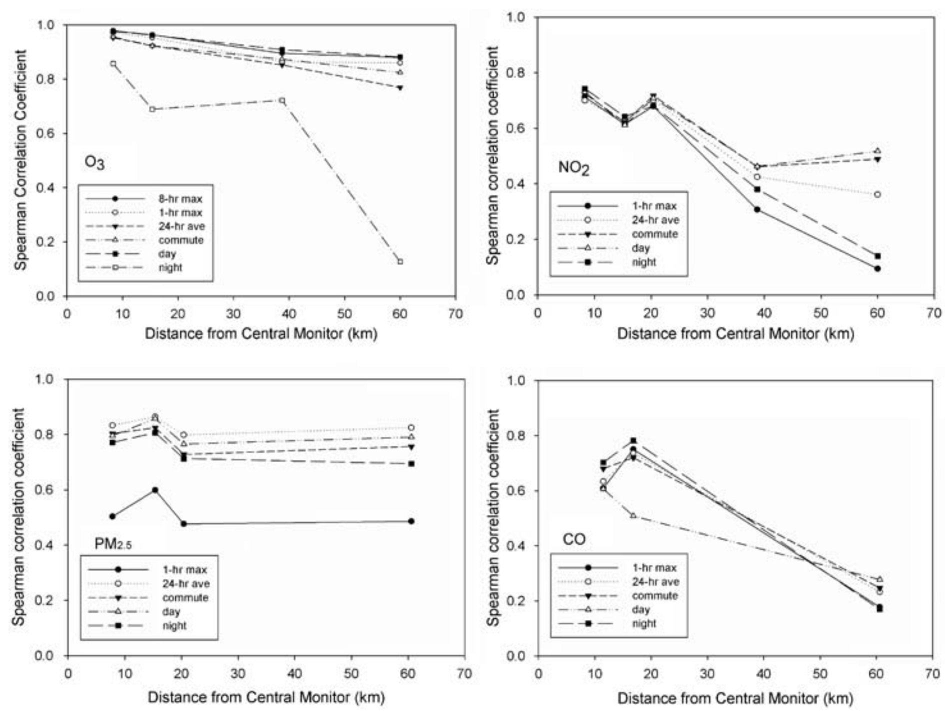

Figure 2. Spatial correlations for O_3, NO_2, $PM_{2.5}$ and CO metrics.

Table 3. Risk ratios, 95% confidence intervals and χ^2-values for associations between lag 1 air pollution metrics and respiratory emergency department visits.

Pollutant	N	Metric	IQR	RR (95% CI) per IQR	Standard unit	RR (95% CI) per std unit	χ^2	P-value
CO (p.p.m.)	3486	1-h maximum	1.40	1.014 (1.009, 1.019)	0.5	1.005 (1.003, 1.007)	30.6	<0.001
		24-h average	0.45	1.015 (1.010, 1.019)	0.5	1.016 (1.011, 1.021)	39.4	<0.001
		Commute (0700–1000 hours, 1600–1900 hours)	0.43	1.007 (1.003, 1.011)	0.5	1.008 (1.003, 1.013)	10.6	0.001
		Day-time (0800–1900)	0.30	1.004 (0.999, 1.008)	0.5	1.006 (0.998, 1.014)	2.0	0.156
		Night-time (2400–0600)	0.62	1.011 (1.008, 1.015)	0.5	1.009 (1.007, 1.012)	46.8	<0.001
NO_2 (p.p.b.)	3635	1-h maximum	23.0	1.011 (1.006, 1.016)	10	1.005 (1.003, 1.007)	22.1	<0.001
		24-h average	13.3	1.012 (1.007, 1.017)	10	1.009 (1.005, 1.013)	21.8	<0.001
		Commute (0700–1000 hours, 1600–1900 hours)	13.8	1.008 (1.003, 1.013)	10	1.006 (1.002, 1.010)	11.1	0.001
		Day-time (0800–1900)	11.5	1.003 (0.998, 1.008)	10	1.002 (0.998, 1.007)	1.2	0.282
		Night-time (2400–0600)	22.2	1.016 (1.011, 1.021)	10	1.007 (1.005, 1.009)	42.4	<0.001
O_3 (p.p.b.)	2883	8-h maximum	28.9	1.020 (1.012, 1.028)	25	1.017 (1.010, 1.024)	23.9	<0.001
		1-h maximum	31.0	1.018 (1.010, 1.025)	25	1.014 (1.008, 1.020)	22.7	<0.001
		24-h average	16.2	1.007 (0.999, 1.015)	25	1.011 (0.999, 1.024)	3.4	0.067
		Commute (0700–1000 hours, 1600–1900 hours)	21.3	1.006 (0.998, 1.015)	25	1.007 (0.998, 1.017)	2.2	0.139
		Day-time (0800–1900)	26.7	1.018 (1.010, 1.026)	25	1.017 (1.009, 1.025)	18.3	<0.001
		Night-time (2400–0600)	19.0	0.991 (0.985, 0.997)	25	0.988 (0.980, 0.996)	8.6	0.003
$PM_{2.5}$ ($\mu g\,m^{-3}$)	1660	1-h maximum	17.4	1.004 (0.999, 1.009)	10	1.002 (1.000, 1.005)	2.9	0.089
		24-h average	10.9	1.004 (0.998, 1.010)	10	1.004 (0.998, 1.010)	1.9	0.171
		Commute (0700–1000 hours, 1600–1900 hours)	11.5	1.005 (0.999, 1.011)	10	1.004 (0.999, 1.009)	2.5	0.113
		Day-time (0800–1900)	10.8	1.003 (0.997, 1.009)	10	1.003 (0.997, 1.008)	0.8	0.380
		Night-time (2400–0600)	13.6	1.004 (0.999, 1.010)	10	1.003 (0.999, 1.007)	2.3	0.133

IQR, interquartile range of pollutant metric.

Table 4. χ^2-values and risk ratios[a] for associations between various lags of each pollutant metric and respiratory emergency department visits.

Pollutant	Metric	χ^2 lag0	χ^2 lag1	χ^2 lag2	χ^2 lag3
CO	1-h maximum	36.8 (1.018)	30.6 (1.014)	27.3 (1.013)	26.8 (1.013)
	24-h average	44.6 (1.019)	39.4 (1.015)	29.4 (1.012)	21.8 (1.011)
	Commute	15.6 (1.011)	10.6 (1.007)	10.1 (1.007)	4.1 (1.004)
	Day-time	4.6 (1.007)	2.0 (1.004)	4.2 (1.005)	0.9 (1.002)
	Night-time	60.6 (1.020)	46.8 (1.011)	23.9 (1.008)	24.7 (1.008)
NO_2	1-h maximum	48.7 (1.020)	22.1 (1.011)	45.9 (1.014)	50.9 (1.015)
	24-h average	50.4 (1.021)	21.8 (1.012)	47.2 (1.016)	49.6 (1.017)
	Commute	34.3 (1.017)	11.1 (1.008)	32.3 (1.013)	23.9 (1.012)
	Day-time	7.4 (1.009)	1.2 (1.003)	9.3 (1.008)	10.2 (1.008)
	Night-time	97.1 (1.027)	42.4 (1.016)	59.1 (1.018)	48.1 (1.016)
O_3	8-h maximum	28.7 (1.027)	23.9 (1.020)	33.2 (1.021)	40.1 (1.022)
	1-h maximum	28.4 (1.026)	22.7 (1.018)	35.1 (1.020)	37.0 (1.020)
	24-h average	3.4 (1.009)	3.4 (1.007)	3.1 (1.007)	9.3 (1.011)
	Commute	1.8 (1.007)	2.2 (1.006)	6.1 (1.009)	14.5 (1.014)
	Day-time	19.5 (1.024)	18.3 (1.018)	23.7 (1.019)	28.3 (1.019)
	Night-time	4.3 (0.993)	8.6 (0.991)	3.6 (0.994)	1.7 (0.996)
$PM_{2.5}$	1-h maximum	1.3 (0.996)	2.9 (1.004)	0.1 (0.999)	2.3 (1.004)
	24-h average	0.9 (0.996)	0.2 (1.004)	1.2 (1.003)	3.4 (1.006)
	Commute	0.7 (0.996)	2.5 (1.005)	1.1 (1.003)	5.7 (1.007)
	Day-time	2.6 (0.993)	0.8 (1.003)	0.4 (1.002)	3.8 (1.006)
	Night-time	0.3 (0.998)	2.3 (1.004)	1.7 (1.004)	1.4 (1.003)

[a] RR per interquartile increase in pollutant metric.

χ^2-values as presented in Table 3. On the basis of the χ^2-values, at shorter lags (0 and 1 days), the nighttime metrics for CO and NO_2 were the strongest predictors of respiratory emergency department visits, whereas at longer lags (2 and 3 days), the χ^2-values were similar among the nighttime, daily 1-h maximum, and 24-h average metrics. This result suggests that strong associations with the lag 1 nighttime metric may reflect associations with a longer lag of pollutant concentrations. For example, the nighttime metric included hours (2400–0600 hours) closest in time to the previous day compared with our other metrics of interest. For CO, NO_2, and $PM_{2.5}$, the nighttime metrics were the most strongly correlated to the previous day's concentrations, regardless of previous daytime metric chosen (Supplementary information Table A).

Comparing Magnitude of Association

Figure 3 displays the risk ratios and 95% confidence intervals scaled to each metric's IQR; correlations between our *a priori* metric (shaded) and alternative metrics are shown on the x-axis. For an IQR increase in each metric, risk ratios ranged from 1.004 to 1.015 for CO, 1.003 to 1.016 for NO_2, 0.991 to 1.020 for O_3, and 1.003 to 1.005 for $PM_{2.5}$. When comparing the magnitudes of association, interpretation differed slightly according to how the regression coefficients were scaled: standard unit or IQR (Table 3). For example, the risk ratio estimate for nighttime NO_2 was highest when effects were scaled according to the IQR, but the 24-h average risk ratio was highest when effects were scaled to the standard unit (10 p.p.b.). For CO, the 24-h average had the highest risk ratio for both scaling approaches, but the daily 1-h maximum was second highest when scaled to its IQR (1.4 p.p.m.), and was lowest when scaled to the standard unit (0.5 p.p.m.). The 1-h maximum, 8-h maximum, and daytime O_3 metrics yielded higher risk ratios than the 24-h average, commute, and nighttime O_3 metrics, regardless of scaling. In Figure 3 and Table 4 we only present the results scaled to each metric's IQR so that risk ratios can be compared for the same relative degree of variability.

In sensitivity analyses using our previous time-series modeling approach (Peel et al., 2005), we observed similar patterns in risk ratios and χ^2-values across the metrics. However, using this approach, the commute metrics for CO and NO_2 were not significantly associated with respiratory emergency room visits.

Discussion

In this time-series analysis, we compared associations between various temporal metrics of daily ambient air pollution levels and respiratory health using a large data set of more than one million respiratory emergency department visits. Our motivation was to explore the implications of choice of pollutant temporal averaging time on health risk estimates within a time-series framework.

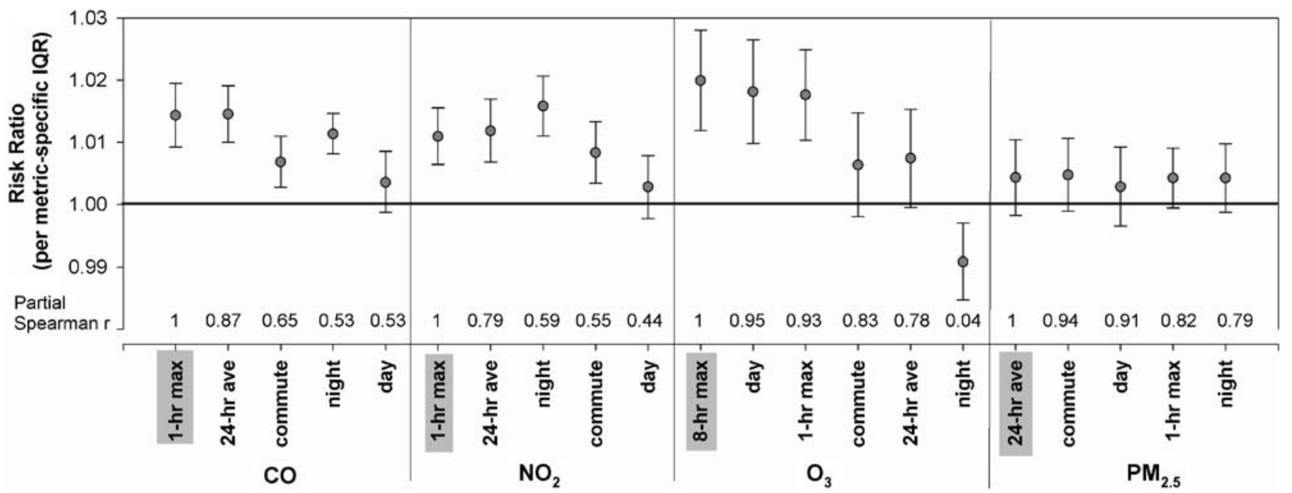

Figure 3. Risk ratios and 95% confidence intervals* for associations between lag 1 pollutant metrics and respiratory emergency department visits. Partial Spearman correlations between *a priori* metrics (shaded in gray) and other pollutant metrics shown above x-axis. *RRs scaled to the interquartile range of each metric.

For a given pollutant, many of the metrics were strongly correlated, and yielded similar magnitude and statistical significance of associations with daily respiratory visits. As expected, pollutant metrics that were less correlated with each other exhibited larger differences in epidemiological associations than did correlated metrics. Differences in epidemiological results between metrics of the same pollutant may be due to (a) differences in biological relevance of the dose for the measured pollutant (for example, peak *vs* average exposures); (b) differences in metric spatial heterogeneity and corresponding representativeness of population exposures (exposure measurement error); (c) differences in correlation with personal exposures due to time-activity patterns (exposure measurement error); (d) differences in representing the true etiological agent (related to surrogacy); (e) differences in representing the relevant lag period of the pollutant (misspecified lag); (f) confounding by other pollutants during certain time-windows (for example, nighttime O_3 with the NO_2 and CO metrics); (g) model misspecification (for example, violation of linearity assumptions); or (h) random variation. Although some of these possible explanations were not directly testable in our study, we discuss them in the context of our findings below.

In general, variability in the observed results reflected pollutant diurnal patterns, with temporal metrics that included peak pollutant hours tending to show the strongest associations with respiratory emergency department visits and metrics capturing hours of low concentrations showing weaker associations. For example, O_3 is formed during the daylight hours and is depleted at night; metrics incorporating the peak afternoon hours of O_3 were correspondingly most strongly associated with our outcome. Conversely, NO_2 is lowest during the daylight hours when it is being more rapidly dispersed and oxidized; during the evening and overnight hours NO_2 is oxidized more slowly and the mixing height decreases, hence the concentrations increase. NO_2 and CO metrics that included the hours of higher concentrations (including the 1-h maximum and nighttime metrics) showed stronger associations than did models using metrics that included concentration minima for these pollutants, despite the typical hours of peak concentration for these pollutants being late evening hours, when people are less likely to be outside.

Differences among the metrics were most pronounced for O_3, wherein three metrics were strongly associated with the outcome (1-h maximum, 8-h maximum, and daytime), two metrics were weakly associated with the outcome (commute and 24-h average), and one metric was inversely associated with the outcome (nighttime). Nighttime O_3 concentrations were negatively correlated with all of the CO, NO_2, and $PM_{2.5}$ metrics, likely due to the depletion of O_3 by reaction with NO; when vehicle emission pollutant concentrations (that is, CO and NOx) are elevated at night, O_3 depletion is high as well. We found that when we controlled for CO and NO_2 concentrations in multi-pollutant models, nighttime O_3 was no longer inversely associated with respiratory emergency department visits. The nighttime concentrations of O_3 may serve as an inverse surrogate of traffic-related pollutants such as NOx. Fresh NO emissions scavenge ozone at night without the reverse process of NO_2 photolysis that leads to O_3 formation. Furthermore, lower mixing heights at night tend to increase NO_x, CO, and $PM_{2.5}$ concentrations. As NO_x emissions are spatially heterogeneous, O_3 scavenging at night is also spatially heterogeneous; in the more populated urban center, the higher NO_x levels result in lower O_3 levels at night. The negative association observed for nighttime O_3 suggests that O_3 is not the only pollutant linked to respiratory outcomes—an example of biological insights gained through assessment of alternative temporal metrics. Lastly, although investigators would be unlikely to choose a nighttime metric of O_3 in an epidemiological study, this finding also argues against the use of a 24-h average metric for O_3, which itself was only weakly associated with the outcome. In our analysis, inclusion of nighttime O_3 concentrations within a 24-h average not only dilutes the relevant concentrations by adding irrelevant hours (that is, bias toward the null) but also includes hours when the relationship between O_3 and respiratory emergency room visits may be negatively confounded by other pollutants.

In a previous study, Bell et al. (2005) compared air quality under seven emission scenarios using 1-h maximum, 8-h maximum, and 24-h average O_3 concentrations to characterize air quality. Rankings of the different emission scenarios differed according to the metric of O_3 chosen, but rankings based on the 1-h maximum and 8-h maximum were more similar to each other than to the rankings based on the 24-h average. In a panel study of asthma symptoms in 25 asthmatic children, associations using the 1-h maximum O_3 metric were similar to those using the 8-h maximum (Delfino et al., 1998). Our O_3 findings are consistent with these previous studies.

From an exposure stand point, our results highlight the potential for certain pollutant metrics to act as surrogates for other pollutant metrics. Our CO and NO_2 findings, for example, indicate that CO or NO_2 could be acting as a surrogate for the other, as shown by the correlations in Table 2 (for example, $r = 0.61$ for 1-h max). Alternatively, CO and NO_2 may be serving as surrogates of another pollutant, namely O_3, as CO and NO_2 metrics incorporating peak concentration hours were shown to be the most strongly associated with the outcome, and more strongly correlated with peak O_3 concentrations than did other CO and NO_2 metrics. Thus, associations between emergency department visits and peak (1-h maximum) NO_2, a precursor of O_3, may be partly confounded by O_3 or vice versa.

Teasing out the effects of each pollutant through the use of multi-pollutant modeling is complicated by the differences in measurement error between the pollutants (Tolbert et al., 2007). Comparative analysis of temporal metrics within and across pollutants may provide an alternative approach for identifying the pollutant more likely to be the etiological agent. For example, NO_2 metrics that were more strongly correlated with 8-h maximum O_3 showed stronger associations with respiratory emergency room visits. This was also true for CO. Furthermore, the CO and NO_2 metrics yielding the strongest associations were the hours when people were least likely to be exposed to ambient air (that is, night hours), and weaker associations were observed when people were more likely to be exposed to ambient air (that is, daytime and commute hours). If NO_2 and CO were the true etiological agents, we might expect to observe associations for metrics incorporating hours when people are more likely to be exposed to ambient air, regardless of the correlation with 8-h maximum O_3. These results suggest that the etiological agent is more likely to be O_3 than CO or NO_2.

Certain metrics may also serve as surrogates of the same pollutant but for a different lag. The nighttime metrics for previous day (lag 1) CO and NO_2 showed some of the strongest associations with respiratory emergency room visits, despite night hours being some of the least likely hours of population exposure to ambient air. Perhaps the nighttime metric at lag 1 is more predictive of respiratory emergency room visits, because it acts as a better surrogate for pollution on earlier days (longer lags). Nighttime NO_2 and CO concentrations were the best surrogates of NO_2 and CO concentrations on the previous day regardless of metric, likely because nighttime hours (2400–0600 hours) were closer in time to the previous day (Supplementary Table A).

The comparison among $PM_{2.5}$ metrics was less informative because we did not observe significant associations with the outcome of interest. However, the consistency of effect estimates across metrics provided reassurance that a strong association would not be missed if analysis was limited to the standard 24-h average metric. In our data, the spatial correlations of the $PM_{2.5}$ metrics were also similar, except for the daily 1-h maximum, which was more spatially heterogeneous. Although in this analysis we did not assess specific chemical components of $PM_{2.5}$, it should be noted that the spatial heterogeneity of $PM_{2.5}$ can vary by the chemical composition of the particles (Wade et al., 2006). Few epidemiological studies have presented results for $PM_{2.5}$ or PM_{10} using a temporal metric other than the 24-h average. We noted reports of three panel studies of asthmatic children investigating the relationship between PM_{10} and asthma symptoms that examined more than one averaging time for PM_{10}. One study found modestly stronger associations using a 24-h average PM_{10} metric compared with a 1-h maximum (Ostro et al., 2001) and two studies showed slightly stronger associations using peak PM_{10} metrics (1-h maximum and 8-h maximum) compared with the 24-h average (Delfino et al., 1998, 2002).

This analysis highlights some of the challenges involved in comparing scaled risk ratios. We presented the risk ratios scaled to each metric's IQR; these risk ratios take into account the range of concentrations for each metric and provide a comparison for the same relative degree of variability. We also presented results for a standard unit (for example, 0.5 p.p.m.) so that results could be compared for the same absolute unit increases in concentration for each pollutant. However, comparisons based on absolute increases in concentration may be misleading in this setting in which daily temporal metrics of the same pollutant are being compared, as a 0.5 p.p.m. increase in daytime CO is a meaningfully greater relative increase compared with a 0.5 p.p.m. increase in 1-h maximum CO, for example. As a consequence of these differences among metrics in concentration variability, the metrics yielding the largest magnitude of effects often differed by the choice of scaling. In this analysis, we preferred the χ^2-values to identify the strongest associations, as χ^2-values are not affected by scaling. Comparing χ^2-values across metrics was appropriate in this study because the sample size was the same for all metrics of a given pollutant.

In this analysis, we focused on respiratory-related emergency department visits and did not present results for cardiovascular disease visits. In our previous work, we found same-day pollution levels (lag 0) to be most strongly associated with cardiovascular emergency room visits (Metzger et al., 2004). Same-day pollution effects can be difficult to compare across temporal metrics because some of the averaging times include hours late in the day, potentially after the bulk of emergency room visits have occurred on a given day. In the present analyses of various metrics of lag 1 pollution, although temporality issues may still have a role (for example, the nighttime metric captured hours at a longer lag than did the daytime metric), all temporal metrics included concentrations before the emergency room visit occurred. These temporality and choice of lag issues are clearly important to the estimation of effects, as recently shown by Lokken et al. (2009).

In summary, we found that epidemiological results were generally similar across different temporal metrics of the same pollutant and would have led to similar conclusions about the relationship between the pollutant and respiratory emergency room visits. Exceptions included the nighttime O_3 metric and the daytime metrics of CO and NO_2. It would be of interest to know how well each of the time-averaged metrics correlate with measured personal exposures; studies in which personal exposures have been measured longitudinally could likely address this question without additional data collection. We found that our *a priori* metrics for CO (1-h maximum), NO_2 (1-h maximum), and O_3 (8-h maximum), based on the National Ambient Air Quality Standards and designed to capture peak concentrations, yielded associations that were as strong or stronger than the other metrics. Our analysis supports the use of these exposure metrics in future studies of ambient air pollution and respiratory health.

Acknowledgements

This work was supported by grants from the Electric Power Research Institute (EP-P27723/C13172), the US Environmental Protection Agency (CR-83407301-0, RD833626, R82921301-0), and the National Institute of Environmental Health Sciences (R01ES11294).

References

Bell M.L., Hobbs B.F., and Ellis H. Metrics matter: conflicting air quality rankings from different indices of air pollution. *J Air Waste Manag Assoc* 2005: 55(1): 97–106.

Delfino R.J., Zeiger R.S., Seltzer J.M., and Street D.H. Symptoms in pediatric asthmatics and air pollution: differences in effects by symptom severity, anti-inflammatory medication use and particulate averaging time. *Environ Health Perspect* 1998: 106(11): 751–761.

Delfino R.J., Zeiger R.S., Seltzer J.M., Street D.H., and McLaren C.E. Association of asthma symptoms with peak particulate air pollution and effect modification by anti-inflammatory medication use. *Environ Health Perspect* 2002: 110(10): A607–A617.

Hansen D.A., Edgerton E., Hartsell B., Jansen J., Burge H., Koutrakis P., Rogers C., Suh H., Chow J., Zielinska B., McMurry P., Mulholland J., Russell A., and Rasmussen R. Air quality measurements for the aerosol research and inhalation epidemiology study. *J Air Waste Manag Assoc* 2006: 56(10): 1445–1458.

Ito K., Thurston G.D., Nadas A., and Lippmann M. Monito-to-monitor temporal correlation of air pollution and weather variables in the North-Central US. *J Expo Sci Environ Epidemiol* 2001: 11(1): 21–32.

Lokken R.P., Wellenius G.A., Coull B.A., Burger M.R., Schlaug G., Suh H.H., and Mittleman M.A. Air pollution and risk of stroke: underestimation of effect due to misclassification of time of event onset. *Epidemiology* 2009: 20(1): 137–142.

Lu Y., and Zeger S.L. On the equivalence of case-crossover and time series methods in environmental epidemiology. *Biostatistics* 2007: 8(2): 337–344.

Metzger K.B., Tolbert P.E., Klein M., Peel J.L., Flanders W.D., Todd K., Mulholland J.A., Ryan P.B., and Frumkin H. Ambient air pollution and cardiovascular emergency department visits. *Epidemiology* 2004: 15(1): 46–56.

Ostro B., Lipsett M., Mann J., Braxton-Owens H., and White M. Air pollution and exacerbation of asthma in African-American children in Los Angeles. *Epidemiology* 2001: 12(2): 200–208.

Peel J.L., Tolbert P.E., Klein M., Metzger K.B., Flanders W.D., Todd K., Mulholland J.A., Ryan P.B., and Frumkin H. Ambient air pollution and respiratory emergency department visits. *Epidemiology* 2005: 16(2): 164–174.

Sarnat J.A., Marmur A., Klein M., Kim E., Russell A.G., Sarnat S.E., Mulholland J.A., Hopke P.K., and Tolbert P.E. Fine particle sources and cardiorespiratory morbidity: an application of chemical mass balance and factor analytical source-apportionment methods. *Environ Health Perspect* 2008: 116(4): 459–466.

Sarnat J.A., Schwartz J., Catalano P.J., and Suh H.H. Gaseous pollutants in particulate matter epidemiology: confounders or surrogates? *Environ Health Perspect* 2001: 109(10): 1053–1061.

Sarnat S.E., Klein M., Sarnat J.A., Mulholland J., Flanders W.D., Waller L.A., Russell A.G., and Tolbert P.E. An examination of exposure measurement error from air pollutant spatial variability in time-series studies. *J Expo Sci Environ Epidemiol*, advance online publication, March 11, 2009; doi:10.1038/jes.2009.10.

Schwartz J. The effects of particulate air pollution on daily deaths: a multi-city case crossover analysis. *Occup Environ Med* 2004: 61(12): 956–961.

Schwartz J. How sensitive is the association between ozone and daily deaths to control for temperature? *Am J Respir Crit Care Med* 2005: 171(6): 627–631.

Sunyer J., Spix C., Quénel P., Ponce-de-León A., Pönka A., Barumandzadeh T., Touloumi G., Bacharova L., Wojtyniak B., Vonk J., Bisanti L., Schwartz J., and Katsouyanni K. Urban air pollution and emergency admissions for asthma in four European cities: the APHEA Project. *Thorax* 1997: 52(9): 760–765.

Tolbert P.E., Klein M., Peel J.L., Sarnat S.E., and Sarnat J.A. Multipollutant modeling issues in a study of ambient air quality and emergency department visits in Atlanta. *J Expo Sci Environ Epidemiol* 2007: 17(Suppl 2): S29–S35.

Van Loy M., Bahadori T., Wyzga R., Hartsell B., and Edgerton E. Aerosol Research and Inhalation Epidemiology Study (ARIES): PM2.5 mass and aerosol component concentrations and sampler intercomparisons. *J Air Waste Manage Assoc* 2000: 50: 1446–1458.

Wade K.S., Mulholland J.A., Marmur A., Russell A.G., Hartsell E.E., Edgerton E., Klein M., Waller L., Peel J.L., and Tolbert P.E. Effect of instrument precision and spatial variability on the assessment of the temporal variation of ambient air pollution in Atlanta, Georgia. *J Air Waste Manage Assoc* 2006: 56: 876–888.

Zanobetti A., and Schwartz J. Air pollution and emergency admissions in Boston, MA. *J Epidemiol Community Health* 2006: 60(10): 890–895.

Zmirou D., Schwartz J., Saez M., Zanobetti A., Wojtyniak B., Touloumi G., Spix C., Ponce de León A., Le Moullec Y., Bacharova L., Schouten J., Pönkä A., and Katsouyanni K. Time-series analysis of air pollution and cause-specific mortality. *Epidemiology* 1998: 9(5): 495–503.

Supplementary Information accompanies the paper on the Journal of Exposure Science and Environmental Epidemiology website (http://www.nature.com/jes)