Statistical Methods for Assessing the Health Risks of Particulate Matter Components

Roger D. Peng

Department Biostatistics
Johns Hopkins Bloomberg School of Public Health

Interface 2008

Old Questions

Total Mass

Okay, but Which Part?

New Questions

Components

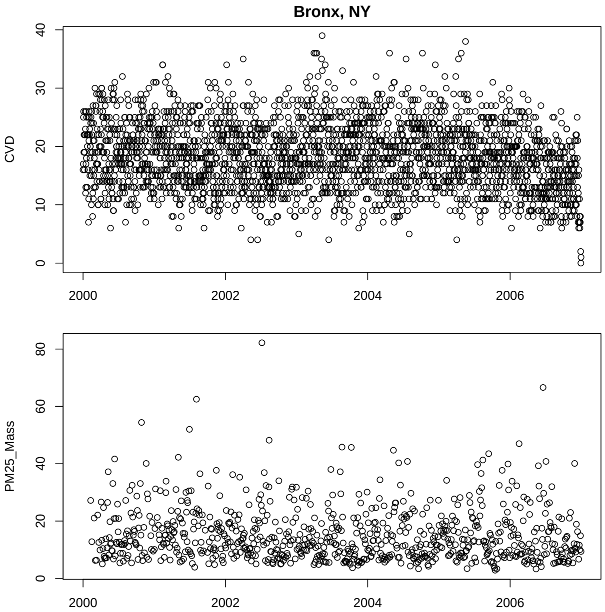
Health Data

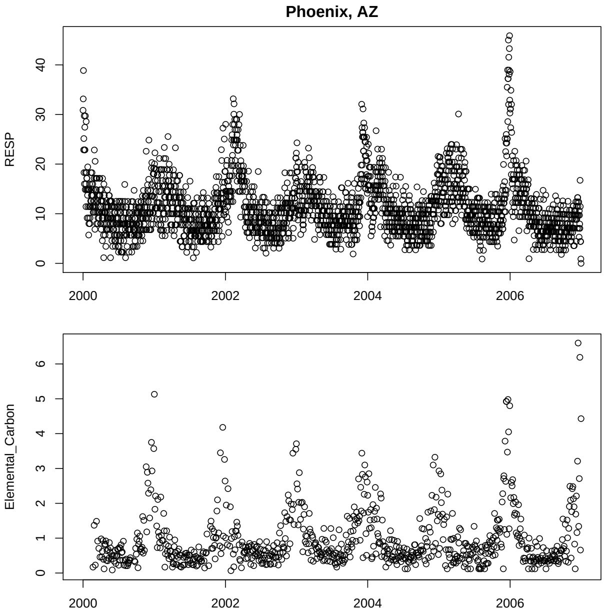
Ambient Pollution Data

Other Data

Assembling National Databases

Total # of Counties	119
Time Period	1999–2006
Study Population	12 million Medicare enrollees
	($\geq$ 65 years of age)
Outcomes	4.3 million CVD admissions
	1.7 million respiratory admissions
Pollutants	Criteria pollutants
	52 Components of $PM_{2.5}$


Bronx, NY
CVD
0
10
20
30
40
2000
2002
2004
2006
PM25_Mass
0
20
40
60
80
2000
2002
2004
2006


Phoenix, AZ
RESP
0
10
20
30
40
2000
2002
2004
2006
Elemental_Carbon
0
1
2
3
4
5
6
2000
2002
2004
2006

Health Outcome Model

Basic model for a single county:

$$\log \mathbb{E}[Y_t] = \beta\, w_t + s(t, \lambda) + \mathbf{z}_t' \gamma$$

- Y_t: count of the number of hospital admissions for a primary diagnosis of a group of ICD-9 codes.
- w_t: ambient concentration from monitoring (e.g. $PM_{2.5}$, coarse PM, chemical components)
- $s(t, \lambda)$: smooth function of time with λ degrees of freedom
- $\mathbf{z}_t$: vector of other variables, i.e. temperature, day of week, humidity

Standard Analysis of Multi-site Time Series Data

Basic model for a single county:

$$\log \mathbb{E}[Y_t^c] = \beta^c \, w_t^c + s^c(t, \lambda) + \mathbf{z}_t^{c\prime} \, \gamma^c$$

We use a Normal approximation, so given $\hat{\beta}^c$ and $\widehat{V}^c$ from each county,

$$\begin{aligned} \hat{\beta}^c \mid \beta^c \quad &\sim \quad \mathcal{N}(\beta^c, \, \widehat{V}^c) \\ \beta^c \mid \mu, \Sigma \quad &\sim \quad \mathcal{N}(\mu, \, \Sigma) \end{aligned}$$

We place diffuse priors on μ and Σ and sample from their posterior distributions.

Fine Particulate Air Pollution and Hospital Admission for Cardiovascular and Respiratory Diseases

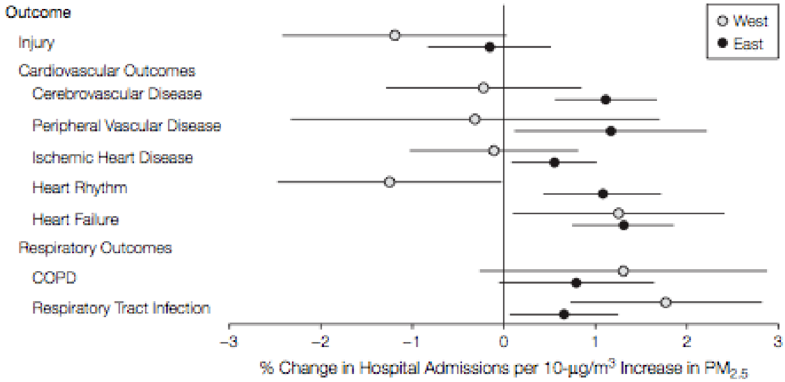
Francesca Dominici, PhD
Roger D. Peng, PhD
Michelle L. Bell, PhD
Luu Pham, MS
Aidan McDermott, PhD
Scott L. Zeger, PhD
Jonathan M. Samet, MD

Context Evidence on the health risks associated with short-term exposure to fine particles (particulate matter ≤2.5 μm in aerodynamic diameter [$PM_{2.5}$]) is limited. Results from the new national monitoring network for $PM_{2.5}$ make possible systematic research on health risks at national and regional scales.

Objectives To estimate risks of cardiovascular and respiratory hospital admissions associated with short-term exposure to $PM_{2.5}$ for Medicare enrollees and to explore heterogeneity of the variation of risks across regions.

Design, Setting, and Participants A national database comprising daily time-series data daily for 1999 through 2002 on hospital admission rates (constructed from

Figure 4. Percentage Change in Hospitalization Rate by Cause per 10-μg/m^3 Increase in $PM_{2.5}$ for the US Eastern and Western Regions for all Outcomes



May 2008, *Journal of the American Medical Association*

Coarse Particulate Matter Air Pollution and Hospital Admissions for Cardiovascular and Respiratory Diseases Among Medicare Patients

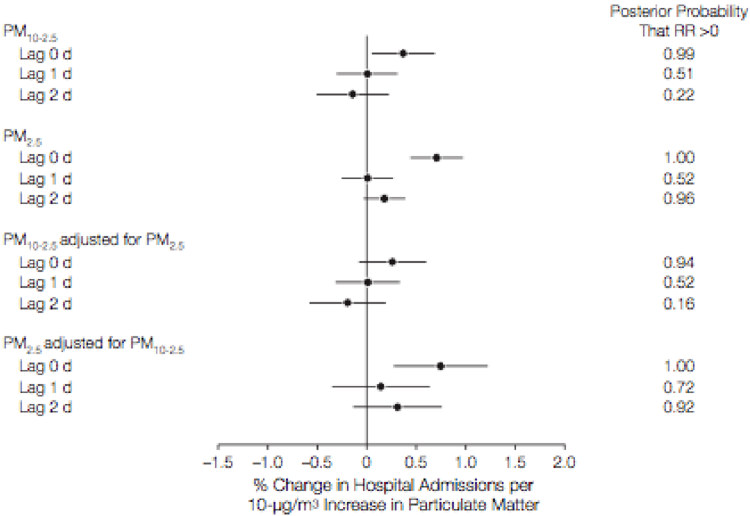
Roger D. Peng, PhD
Howard H. Chang, BS
Michelle L. Bell, PhD
Aidan McDermott, PhD
Scott L. Zeger, PhD
Jonathan M. Samet, MD
Francesca Dominici, PhD

Context Health risks of fine particulate matter of 2.5 μm or less in aerodynamic diameter ($PM_{2.5}$) have been studied extensively over the last decade. Evidence concerning the health risks of the coarse fraction of greater than 2.5 μm and 10 μm or less in aerodynamic diameter ($PM_{10-2.5}$) is limited.

Objective To estimate risk of hospital admissions for cardiovascular and respiratory diseases associated with $PM_{10-2.5}$ exposure, controlling for $PM_{2.5}$.

Design, Setting, and Participants Using a database assembled for 108 US counties with daily cardiovascular and respiratory disease admission rates, temperature and dew-point temperature, and $PM_{10-2.5}$ and $PM_{2.5}$ concentrations were calculated with

Figure 2. Percentage Change in Emergency Hospital Admissions Rate for Cardiovascular Diseases per a 10-μg/m³ Increase in Particulate Matter



Study sees threat from big-particle pollutants

Tue May 13, 2008 4:06pm EDT

Email | Print | Share | Reprints | Single Page | Recommend (0)   [-] Text [+]

By Andrew Stern

Study sees threat from big-particle pollutants

Tue May 13, 2008 4:06pm EDT

Email | Print | Share | Reprints | Single Page | Recommend (0) [-] Text [+]

By Andrew Stern

Go Ahead -- Take a Deep Breath!

(Ivanhoe Newswire) -- Very little research has been done on the health effects of
particulate air pollution on cardiovascular disease. But a new study by researcher
Hopkins Bloomberg School of Public Health found risks from exposure to these k
pollutants was not significantly linked with increased hospital admissions.

Study sees threat from big-particle pollutants

Tue May 13, 2008 4:06pm EDT

Email | Print | Share | Reprints | Single Page | Recommend (0)   [-] Text [+]

By Andrew Stern

Go Ahead -- Take a Deep Breath!

(Ivanhoe Newswire) -- Very little research has been done on the health effects of coarse particulate air pollution on cardiovascular disease. But a new study by researchers at Johns Hopkins Bloomberg School of Public Health found risks from exposure to these kinds of pollutants was not significantly linked with increased hospital admissions.

Health News

Not All Air Pollution Harms Equally

FONT SIZE A A A

Study Shows Larger Pollution Particles May Not Play Large Role in Hospital Admissions

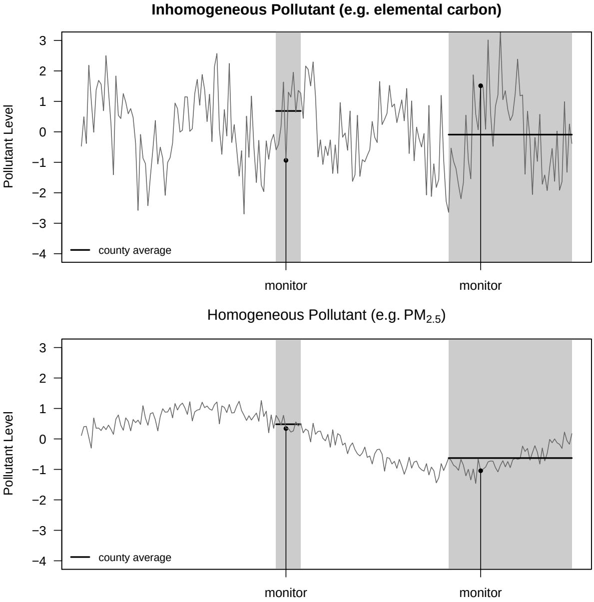
By Kelley Colihan
WebMD Medical News

Reviewed by Louise Chang, MD

Preliminary Work: Adjusting for Spatial Heterogeneity

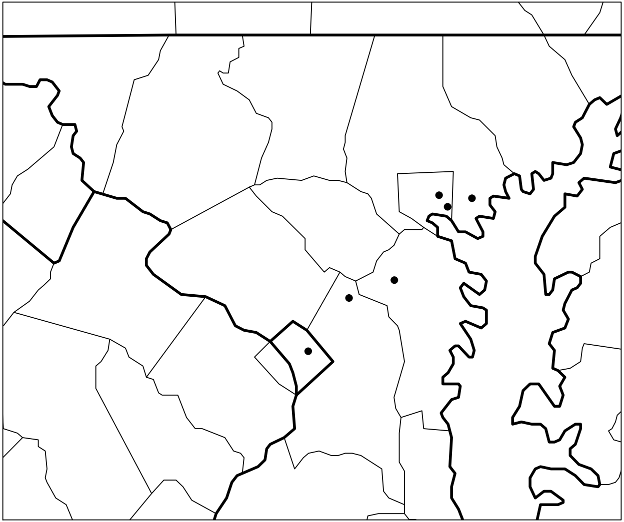
Coarse PM

$PM_{2.5}$ Components

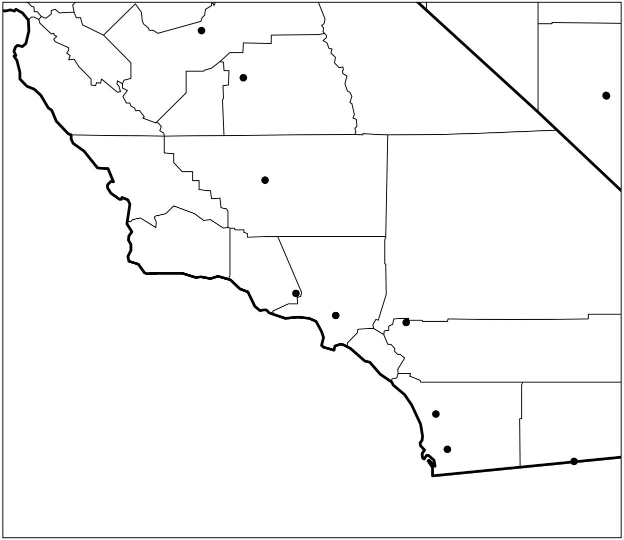


Inhomogeneous Pollutant (e.g. elemental carbon)
Pollutant Level
3
2
1
0
−1
−2
−3
−4
county average
monitor
monitor
Homogeneous Pollutant (e.g. PM2.5)
Pollutant Level
3
2
1
0
−1
−2
−3
−4
county average
monitor
monitor

STN Monitors in Maryland/Washington, D.C.



STN Monitors in Southern California



Health Outcome Model

Basic model for a single county:

$$\log \mathbb{E}[Y_t] = \beta\, w_t + s(t, \lambda) + \mathbf{z}_t'\, \gamma$$

- Y_t: count of the number of hospital admissions for a primary diagnosis of a group of ICD-9 codes.
- w_t: ambient concentration from monitoring (e.g. chemical components, coarse PM)
- $s(t, \lambda)$: smooth function of time with λ degrees of freedom
- $\mathbf{z}_t$: vector of other variables, i.e. temperature, day of week, humidity

Error Model

Health outcome model:

$$\log \mathbb{E}[Y_t] = \beta\, w_t + s(t, \lambda) + \mathbf{z}_t'\, \gamma$$

We measure w_t, the monitor value (typically 1 monitor per county).

We want $\bar{x}_t$, the county-wide average of the pollutant (i.e. if we had infinite monitors)

The error model is then

$$w_t = \bar{x}_t + u_t$$

where $\mathbb{E}[u_t] = 0$ and $\text{Var}(u_t) = \sigma_u^2$.

We can calculate $\hat{\beta}$ using w_t but we want an estimate of β that we would have gotten if we had used $\bar{x}_t$

Underlying Spatial Model

We assume that the pollutant follows an underlying spatial model from which we draw observations at the monitor location.

$$z(s,t) = \mu(s,t) + \varepsilon(s,t)$$

so that

$$\begin{aligned} w_t &= z(s_m, t) \;\; \text{[evaluated at monitor location } m\text{]} \\ \bar{x}_t &= \frac{1}{|A|} \int z(s,t)\, ds \end{aligned}$$

where A is the county boundary. We assume a Gaussian process for ε, i.e.

$$\varepsilon(\cdot, t) \sim \text{GP}(0, \tau^2 \rho(\cdot, \cdot; \theta))$$

with correlation function ρ.

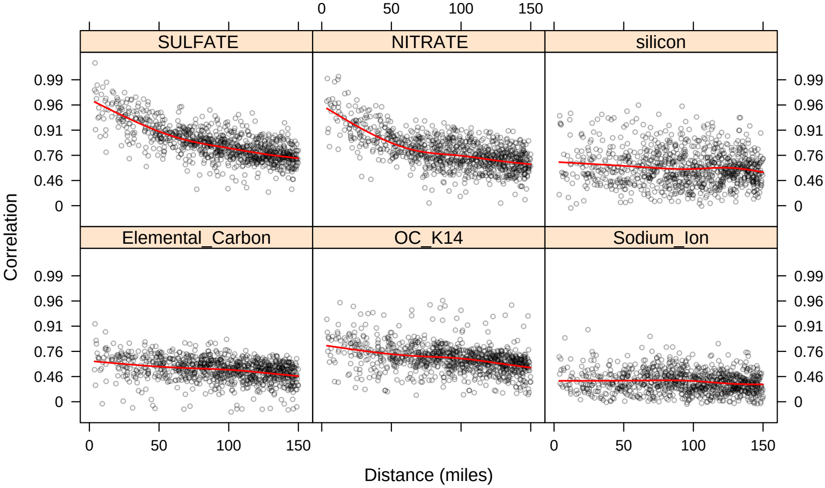
Underlying Spatial Model

Using the spatial model and assuming a Gaussian process, we can calculate

$$\sigma_u^2 = \tau^2 \left(1 - \frac{2}{|A|} \int \rho(s_m, s; \theta)\, ds + \frac{1}{|A|^2} \iint \rho(s, s'; \theta)\, ds'\, ds \right)$$

which is the error variance due to spatial inhomogeneity. All we need is an estimate of ρ.

Correlograms for $PM_{2.5}$ Components



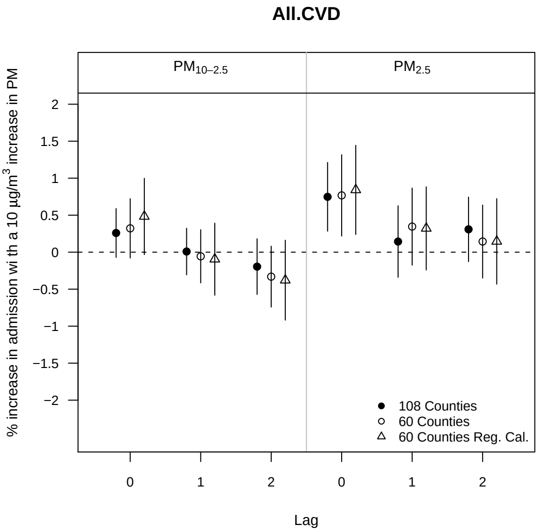
Regression Calibration

Given our spatial model the monitor value w_t and the county average $\bar{x}_t$ have a joint Normal distribution.

If we have and estimate for ρ, we can approximate the distribution of $\bar{x}_t \mid w_t, \mathbf{z}_t$ and obtain the calibration function $\mathbb{E}[\bar{x}_t \mid w_t, \mathbf{z}_t]$.

Replacing w_t with $\mathbb{E}[\bar{x}_t \mid w_t, \mathbf{z}_t]$ would give us the regression calibrated estimate of β.

Regression Calibration and Coarse PM



Simulation Extrapolation (SIMEX)

Given an estimate for σ_u^2, we can simulate pseudovalues

$$w_t^\star = w_t + \sqrt{\alpha} u_t$$

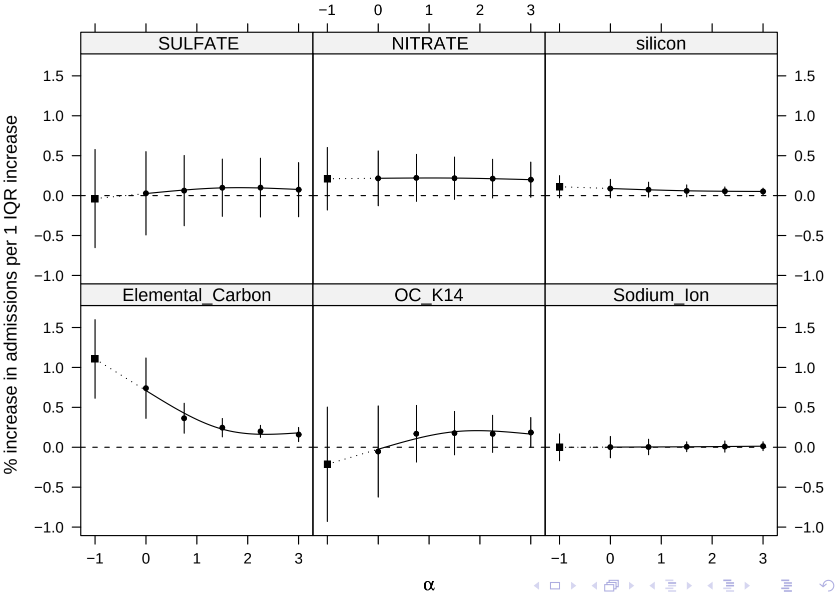
where $u_t \sim \mathcal{N}(0, \sigma_u^2)$. Using the pseudovalues in place of the observed values we can fit our health outcome model and calculate $\hat{\beta}_\alpha$.

Note that

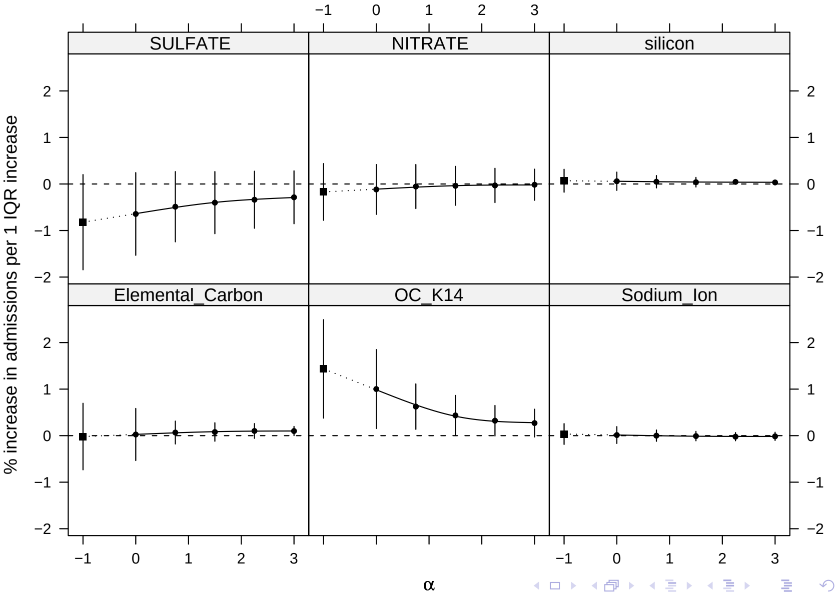
$$\mathrm{Var}(w_t^\star \mid \bar{x}_t) = (1 + \alpha)\sigma_u^2$$

We would like to simulate data for $\alpha = -1$, but we can't so we try values $\alpha = 1, 2, 3, \ldots$ and then extrapolate the $\hat{\beta}_\alpha$s backwards.

SIMEX Results for CVD and Components



SIMEX Results for RESP and Components



Summary

- Examination of components of particulate matter is an important task but introduces new problems
- Spatial modeling of component data can be used to estimate measurement error due to spatial heterogeneity across counties
- Regression calibration and SIMEX both produce estimates that are higher than original risk estimates (30%–100% higher), suggesting that there is some bias
- Current approach is piecewise ad hoc—more work can be done to address uncertainty; perhaps a Bayesian approach to integrate spatial model with the health outcome model

Future Work

- Our pollution data are spatially sparse, but our health data are dense
- Population weighted exposures in counties with > 1 monitor
- No reason to use political boundaries; can use ZIP codes directly surrounding a monitor

Acknowledgements

Collaborators

- Johns Hopkins University: Francesca Dominici, Howard Chang, Aidan McDermott, Scott Zeger, Jon Samet, Alison Geyh
- Yale University: Michelle Bell, Keita Ebisu

Funding

- NIEHS
- EPA Particulate Matter Research Center
- Health Effects Institute