

Analysis of Cardiac Birth Defects and Air Pollution in Texas

Laura Boehm

North Carolina State University, Department of Statistics

Research Questions

- Is Carbon Monoxide (CO) pollution exposure associated with increased risk of Birth Defect?
- Is there a particular window of increased sensitivity to pollution during or before pregnancy?

Previous Research

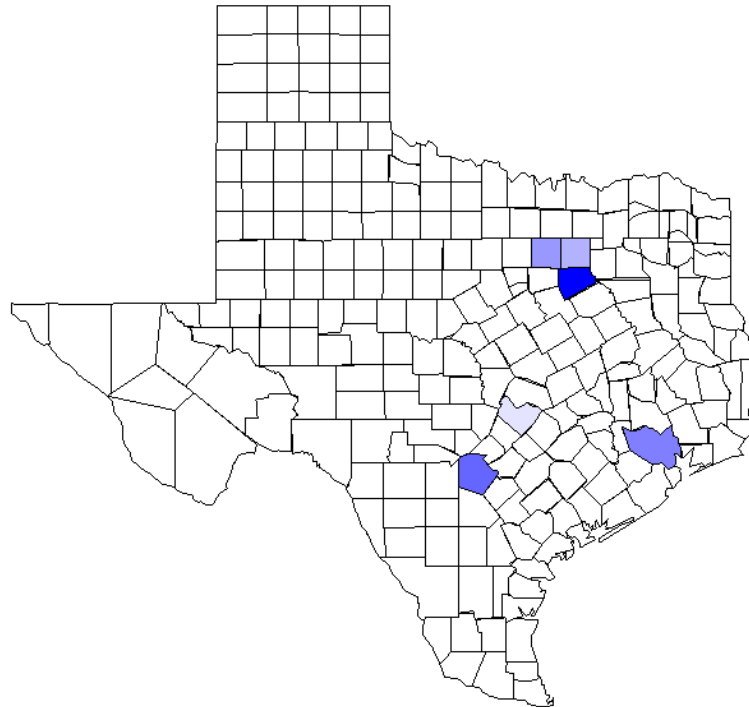
- Gilboa et al. 2005 compared quartiles of pollution exposure for weeks 3-8 of pregnancy
 - Found positive link between CO and tetralogy of Fallot, and between SO and Ventricular Septal Defects (VSD)
 - Inverse relationship between CO and Atrial Septal Defects and Ozone and VSD.
- Gilboa et al. mostly confirms southern California study by Ritz et al.

Data

- Matched Case Control study in 6 Texas counties, 1997-2000
- 9 cardiac birth defects
- Daily Ozone and CO (nearest monitor and County average), and Temperature for 50 days prior and 85 after “conception”
- Covariates: Maternal age, race/ethnicity, years education, alcohol/tobacco use, Illness status, plurality, parity, gravidity, sex of child, season of birth
- 4091 births observed, 3936 with no missing covariates. 2274 cases; 1683 unique individuals with defects.

Data Locations

Counties Included in Study: Mean CO Concentration



Research Plan

- Develop a model framework in which data can “tell” us an exposure window of elevated risk rather than doing hypothesis tests on any given number of windows.
- Somewhat uncommon spatial problem: many observations at the same locations

Variable Selection on Windows

- Create windows and then use Bayesian Model selection techniques to select which windows are significant.
- Ten Gaussian windows with means at day (μ) 7, 21, 35, ... and standard deviation (σ) 7. *Poll* is vector of daily pollution

$$w_{tk}^* = e^{\frac{-1}{2\sigma^2}(t-\mu_k)^2}, \quad w_{tk} = \frac{w_{tk}^*}{\sum_{t=1}^{135} w_{tk}^*}, \quad z_{ik} = \sum_{t=1}^{135} w_{tk} \text{Poll}_{it}$$

- In Matrices: $\mathbf{Z} = \mathbf{POLL} \cdot \mathbf{W}$ so coefficients on \mathbf{Z} , α , can be transformed into coefficients, β , on \mathbf{POLL} , by $\beta = \mathbf{W} \cdot \alpha$

Logistic regression with SVSS

$$Y_i \sim \text{Bin}(1, \pi_i)$$

$$\text{logit}\pi_i = x_i^T \gamma + \sum_{k=1}^{10} \alpha_k z_{ik}$$

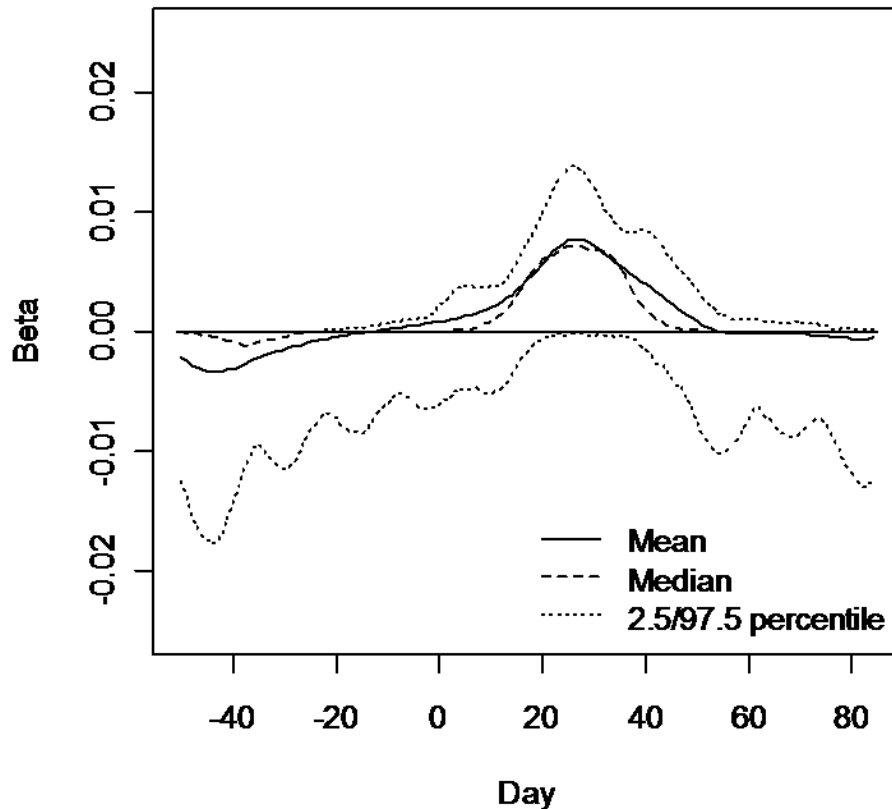
$$g_k \sim \text{Bernoulli}(p_k)$$

$$\alpha_k \sim \text{Normal}(\mu, \tau^2 + (1 - g_k)\tau_2^2)$$

Covariates included are Maternal Age, Education, Race/Ethnicity (Hispanic White, Non Hispanic White, Non Hispanic Black and Other), and Illness status, and Sex of the child, and Plurality.

Results with no Spatial Random Effects

Avg Daily CO, no Spatial RE



If CO exposure increased one standard deviation for the entire study period, odds of VSD would increase 13% (1.13 times)

Credible interval: (.39 , 1.52 times)

If CO exposure increased one standard deviation for weeks 2-6 of pregnancy, odds of VSD would increase 18% (1.18 times)

Credible interval (.97, 1.33 times)

What About Spatial Correlation?

- Overdispersion in the Binomial model due to county/regional effects
- Account for the fact that counties closer together will be more highly correlated in pollution and other unmeasured covariates (including environmental exposures)

Spatial Random Effects with CAR prior

$$Y_i \sim \text{Bin}(1, \pi_i)$$

$$\text{logit}\pi_i = x_i^T \gamma + \sum_{k=1}^{10} \alpha_k z_{ik}$$

$$g_k \sim \text{Bernoulli}(p_k)$$

$$\alpha_k \sim \text{Normal}(\mu, \tau^2 + (1 - g_k)\tau_2^2)$$

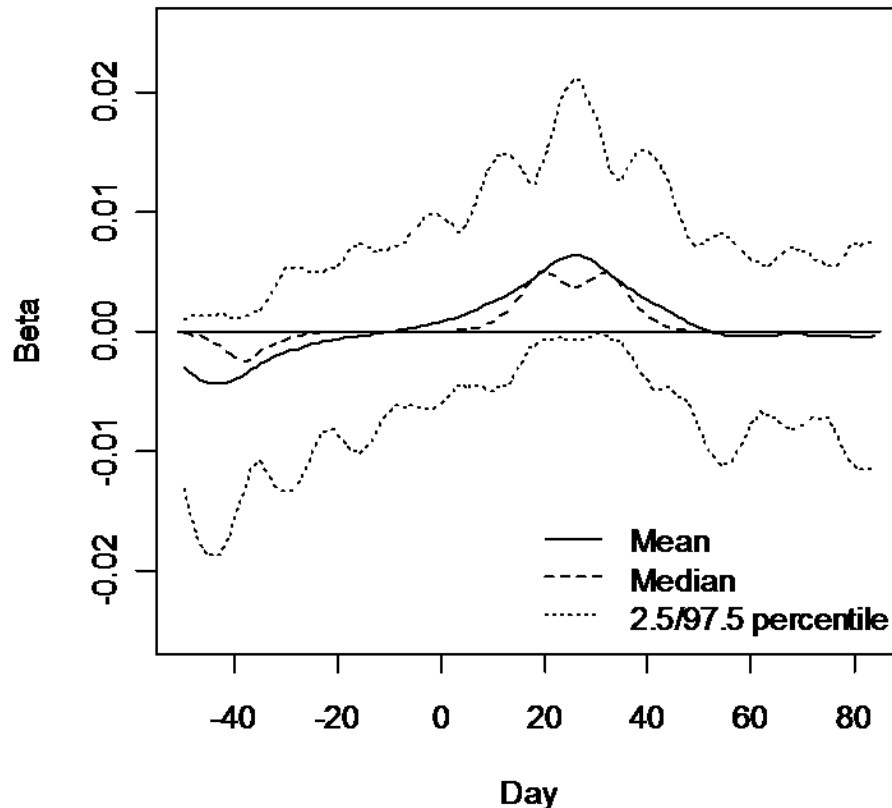
Same SSVS Weighted window model as before.

$$\phi_s | \phi_{r \neq s} \sim N \left(\frac{\sum_{r \neq s} c_{rs} \phi_r}{\sum_{r \neq s} c_{rs}}, \frac{1}{\tau \sum_{r \neq s} c_{rs}} \right)$$

Same covariates included : Maternal Age, Education, Race/Ethnicity , and Illness status, and Sex of the child, Plurality.

Results with Spatial Random Effects

Avg Daily CO, with Spatial RE



If CO exposure increased one standard deviation for the entire study period, odds of VSD would increase 8% (1.08 times)

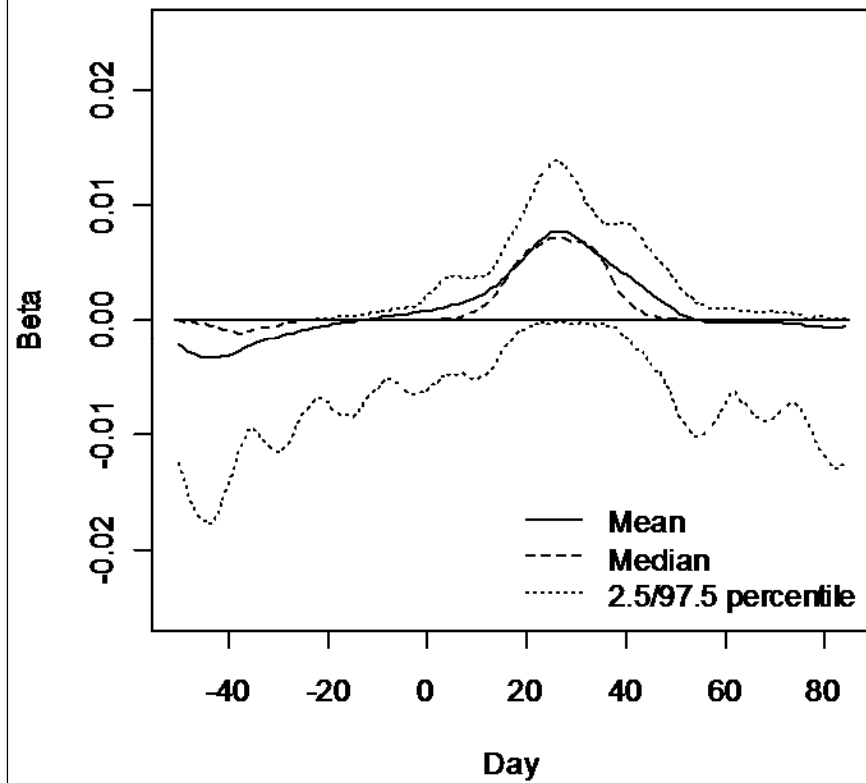
Credible interval: (.36 , 3.2 times).

If CO exposure increased one standard deviation for weeks 2-6 of pregnancy, odds of VSD would increase 14% (1.14 times)

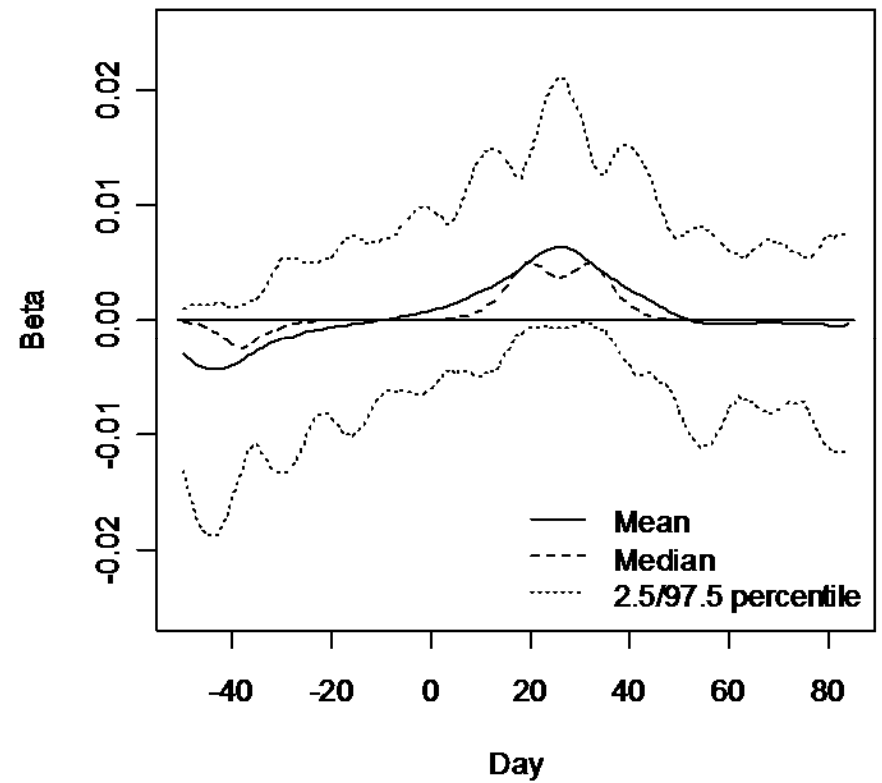
Credible interval (.95, 1.58 times)

Results Comparison

Avg Daily CO, no Spatial RE



Avg Daily CO, with Spatial RE



Other things to explore

- Modeling residuals (remove Seasonal Trend)
- Incorporating temperature
- Multiple defects, multiple pollutants

Thank you!

Questions?

Resources

- George, E. And McCulloch, R. (1996) Stochastic search variable selection. Ch 12 in : Gilks, W. , Richardson, S. , and Spiegelhalter, D. (1996) Markov Chain Monte Carlo in practice: Interdisciplinary statistics. Chapman Hall.
- Gilboa, S. , Mendola, P., et al. (2005) Relation between ambient air quality and selected birth defects, seven county study, Texas, 1997-2000. *American Journal of Epidemiology* 162, 238-252.
- Waller, L. and Gotway, C. (2004) Applied spatial statistics for public health data. John Wiley & Sons.