Modelling Effects of Air Pollution on Acute Asthma Outcomes in Chicago

Paul Rathouz, PhD
prathouz@uchicago.edu
Department of Health Studies
and
Center for Integrating Statistical and Environmental Sciences (CISES)
University of Chicago

March 24, 2004

This research is funded by
U.S. EPA - Science To Achieve Results (STAR) Program
Grant # R-82940201-0

Although the research described herein has been funded wholly or in part by the United States Environmental Protection Agency through STAR Cooperative Agreement # R-82940201-0 to the University of Chicago, it has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.

Project Questions and Objectives

Objective 1:

To model the spatio-temporal behavior of ozone and particulate matter in Chicago

Objective 2:

To develop and apply statistical models to assess the short-term respiratory health effects of daily fluctuations in air pollution in Chicago over four summers in the 1990s.

Environmental health question:

Given the relatively clean urban environment of present-day Chicago, what role does degraded air quality play in the incidence of acute asthma events?

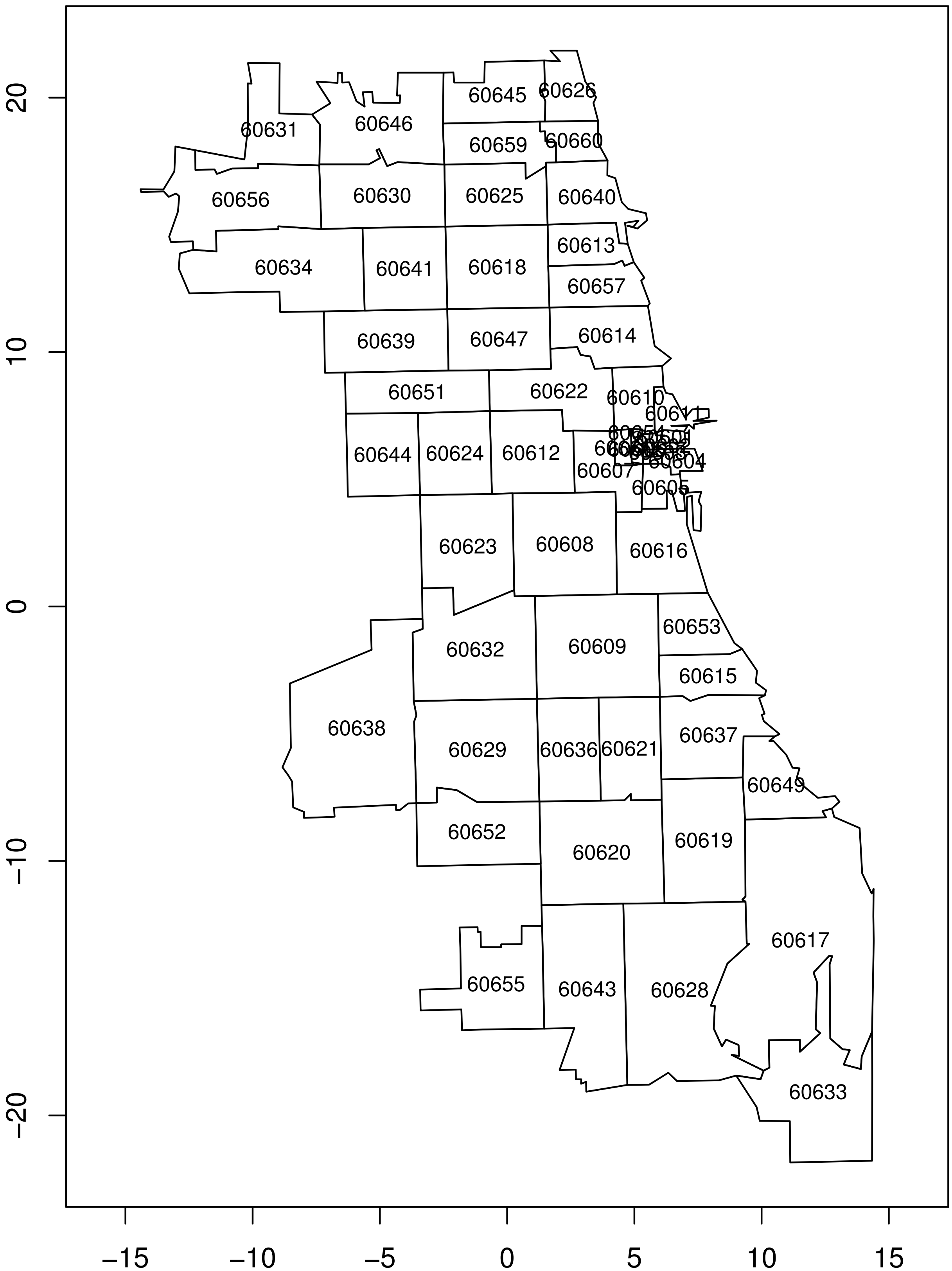
Statistical question:

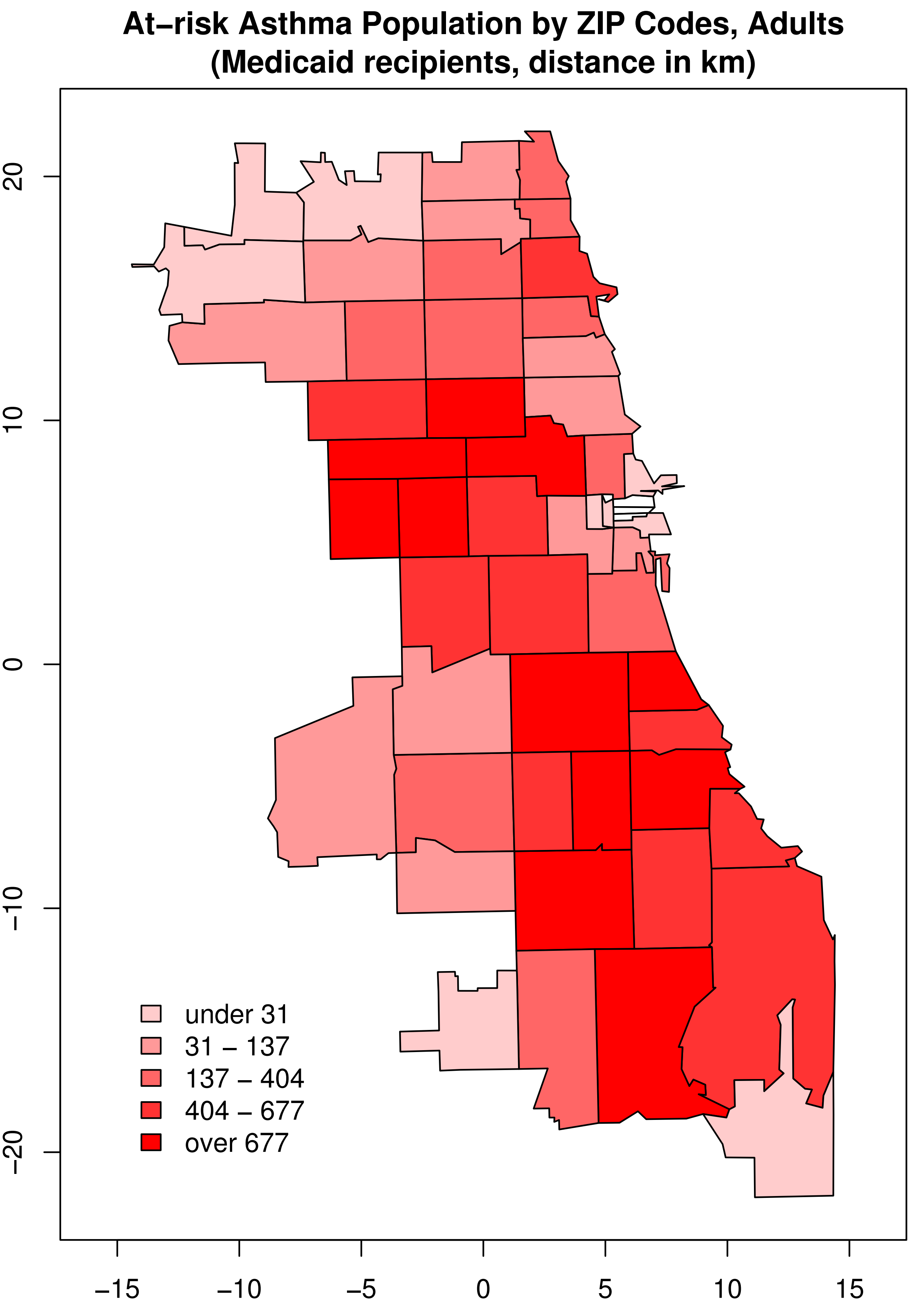
How can one construct models to account adequately for **space-time variability** across the urban area and to exploit **longitudinal data** as they relate to the incidence of asthma?

Chicago Asthma Medicaid Database

- IDPA $200,000$ people on Medicaid in Chicago
- Medicaid (IDPA) claims data
  - "summers" 1995–1998
  - April 1 through October 31
  - Avoids flu season and low ozone season
- "Asthma cohort" $\longleftarrow$ at least one asthma claim during study period
  - About 20,000 adult (15,000 children) asthma patients in our data set
  - Mostly poor, mostly women (adults)
  - Most on Medicaid for full three years
- Event types and dates:
  - Hospitalizations and ED visits
  - Short-term beta-agonists (Albuterol) ($\sim 135,000$ adult; $\sim 58,000$ child events)
- Subject ID, ZIP-code of residence, gender, age

ZIP Code Boundaries




At−risk Asthma Population by ZIP Codes, Adults
(Medicaid recipients, distance in km)
20
10
0
−10
−20
−15
−10
−5
0
5
10
15
under 31
31 – 137
137 – 404
404 – 677
over 677

Chicago Air Pollution Database

- Ozone: 11 monitors
  - Hourly ozone
  - Max-8-hr average ozone
  - Ordinary spatial Kriging assuming isotropy
  - Matérn covariance function
  - Predicted values of ozone (and se) at each ZIP centroid
  - Spatially resolved exposure measure!
- PM10: 18 monitors
  - Observations spaced 6 days apart
  - Monitor and day effect ANOVA model to predict PM10 for any given monitor on any given day
  - Daily average based on the fitted values

Chicago Weather and Pollen Databases

- Weather: Average of O'Hare and Midway:
  - temperature, humidity
  - visibility, wind, precipitation
- Pollen:
  - combined trees, grasses, weeds
  - smoothing used for imputation of missing values

Major Aspects of Project

1. Space-time ozone modeling across urban area
2. **Validation of Beta-agonist as Asthma Outcome**
3. **Mapping asthma outcomes aggregated by ZIP-code**
4. Aggregate data time series analysis of asthma outcomes
5. Spatio-longitudinal analysis of asthma outcomes

Validation of Beta-Agonist (BA) as an Asthma Outcome

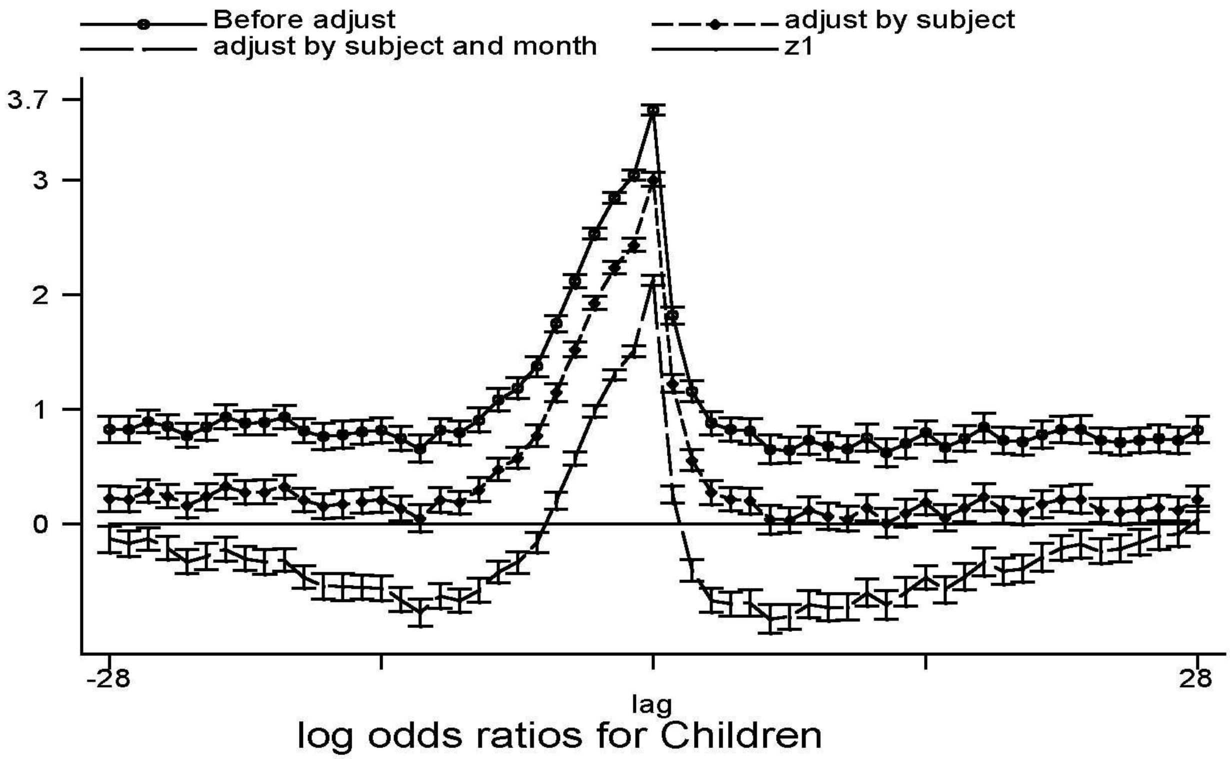
- Why use BA (Albuterol)?
  - often-used non-steroid bronchodilators
  - specific to asthma
  - running prescriptions, quick use
  - large counts

- Q: Is using BA prescriptions as a marker for asthma outcome reasonable?

  $\longrightarrow$ Examine association between BA and hospital admits / ED visits (more traditional outcomes), **lagged** by one or more days

- Odds ratio (OR)
  - crude OR
  - subject-adjusted OR
  - subject- and time-adjusted OR

Association (log OR) between BA and ED/Hospital Admits



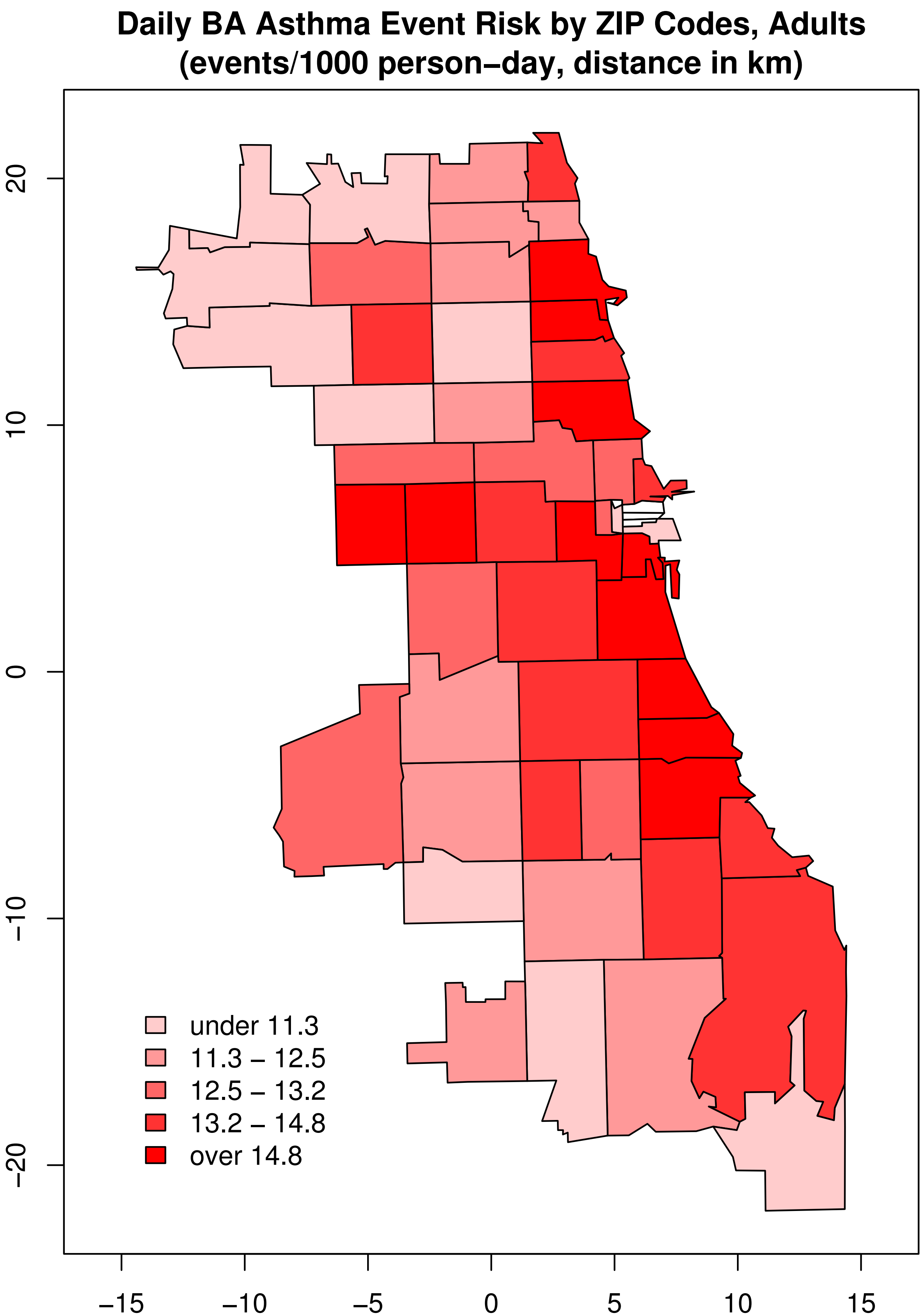
(Note: positive lag = BA before ED)

Problems with crude OR:

...Artifactually high association:

- some subjects in poorer respiratory health
- varying periods of respiratory health, by subject
- does not capture "short-term" association

Mapping asthma outcomes aggregated by ZIP-code



Daily BA Asthma Event Risk by ZIP Codes, Adults
(events/1000 person−day, distance in km)
20
10
0
−10
−20
−15
−10
−5
0
5
10
15
under 11.3
11.3 − 12.5
12.5 − 13.2
13.2 − 14.8
over 14.8

Statistical Problem

Data: observed, **average** (over space and time)
disease risk within ZIP code (**areal unit**)

Target: **predicted** disease risk (given data)
point-by-point

Why model variation in disease risk ?

- description of disease incidence

  . . . spatial interpolation

  $\longrightarrow$ disease risk **mapping**

- place spatial epidemiologic studies in context by displaying background risk

- interpret (variation in) disease risk at **point** versus **area** level

  - smooth variation across areal units
  - account for (differences in) sampling variability across areal units
  - borrow information across units

- understand areal data aggregation

- improve statistical efficiency in spatial regression models

Statistical Approach: Model Risk Aggregation

- $R(x)$ is **log asthma-risk** at location x
- Then . . .

$$Y_i \sim \text{Poisson}\left(\text{person-time} \times e^{\bar{R}_i}\right)$$

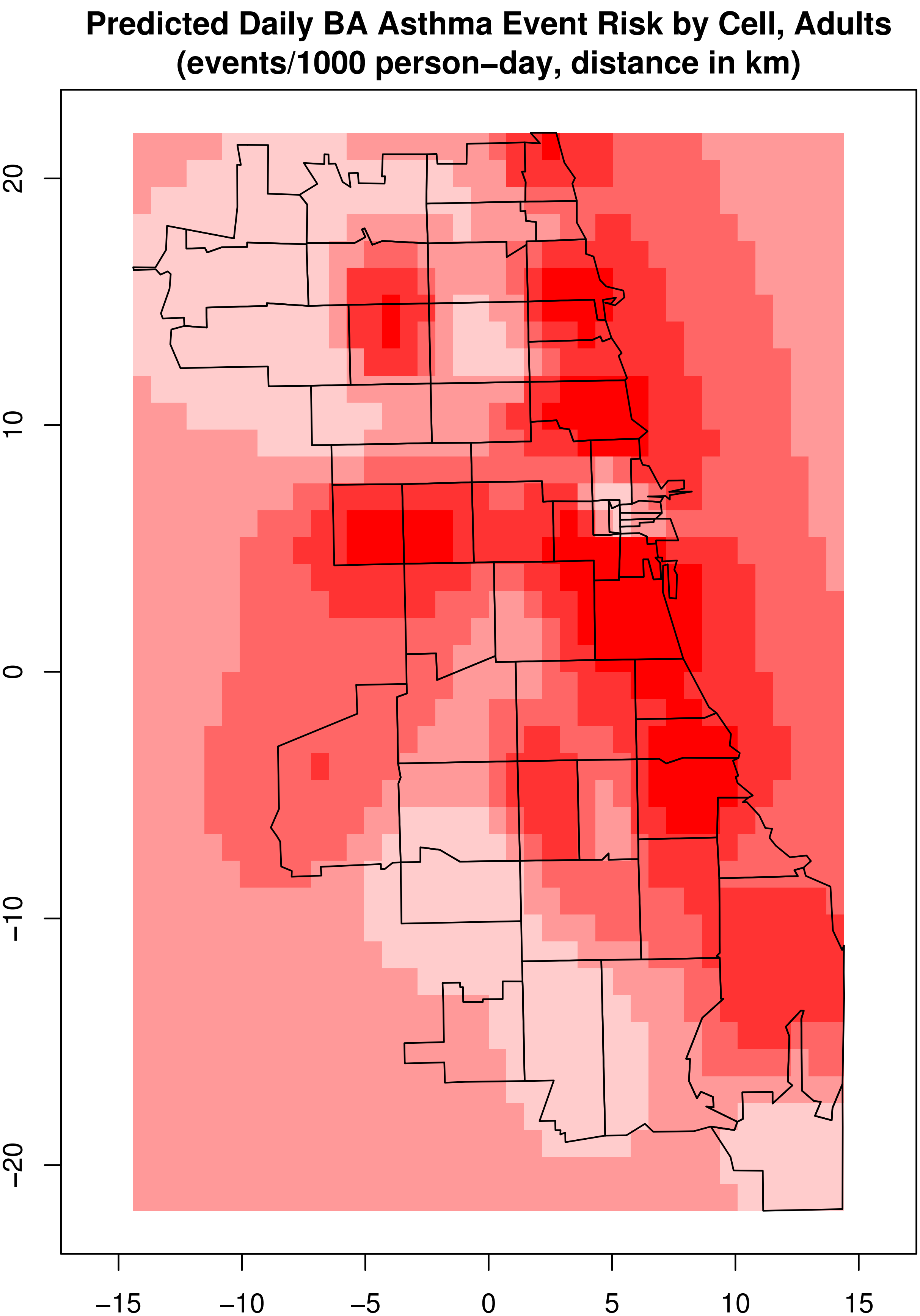
  where $\bar{R}_i = \text{ave}\{R(x)\} =$ **average log-risk** over ZIP i

- Converts **point-scale** problem in $R(\cdot)$ to **area-scale** problem in $\bar{R}_i$
- An estimate of $\bar{R}_i$ is

$$\widehat{\bar{\bar{R}}}_i = \log\left(\frac{\#\ \text{events}}{\text{person-time}}\right),$$

  the observed log risk in ZIP i

- $\widehat{\bar{\bar{R}}}_i$'s are used to **predict** $R(x)$ at every x


Predicted Daily BA Asthma Event Risk by Cell, Adults
(events/1000 person−day, distance in km)
20
10
0
−10
−20
−15
−10
−5
0
5
10
15

Major Aspects of Project

1. **Space-time ozone modeling across urban area**
   - new statistical space-time models
   - exploit physical models such as CMAQ
   - improve air pollution exposure measurement
2. Validation of Beta-agonist as Asthma Outcome
3. Mapping asthma outcomes aggregated by ZIP-code
4. **Aggregate data time series analysis of asthma outcomes**
5. **Spatio-longitudinal analysis of asthma outcomes**

CISES Ozone-Asthma Co-Investigators

Vanja Dukic (Health Studies)

Dana Draghicescu (CISES, Research Associate)

John Frederick (Geophysical Sciences)

Edward Naureckas (Medicine)

Alexis Zubrow (CISES, Programmer/Analyst)

Xiaoming Bao (Statistics, MS Student)

Longitudinal Analysis of Asthma Outcomes

Data:

$Y_{ijt} = I($event, person j, ZIP i, day $t)$

$i =$ ZIP; $j =$ person in ZIP; $t =$day

$x_{ij} =$ (**unobserved**) location

Hazard model:

$$\Pr(Y_{ijt} = 1) = e^{R(x_{ij})+S_{ijk}+\beta' z_{ijt}}$$

where:

- $R(x)$ is baseline **log-asthma risk** at x
- $S_{ijk} =$ all unobserved factors for person (i, j) in time "window" k
- $z_{ijt} =$ covariates for person (i, j), day t

Confounding control and efficiency:

- $\widehat{\beta}$ adjusted for spatially and within-person slowly-varying factors
- $\widehat{\beta}$ has improved statistical efficiency

Model fit: Conditional logistic regression

Longitudinal Data Analysis
Preliminary Model Fit
Adults

Odds ratios for Risk of BA Prescription Fill

Conditional Logit Model

	Odds ratio	St. Err.	Z
ozone (20ppb)	1.005	.007	0.81
pm10 (15mg/m3)	.999	.004	-0.22
log(pollen) (std)	1.009	.004	2.11
temp (5F)	.984	.003	-5.20
rel hum (10%)	.997	.003	-0.95

Average Odds Ratios across 51 ZIP Codes

	Odds ratio	St. Err.	Z
ozone (20ppb)	1.032	.022	1.47