Sources of air pollution and their impacts on human health in multicity studies

Jenna R. Krall

Assistant Professor
Department of Global and Community Health
George Mason University

Background

- Ambient air pollution is a temporally and spatially varying mixture
  - Gases (ozone, carbon monoxide)
  - Particulate matter (PM): size distributions PM_{10}, $PM_{2.5}$
  - PM constituents: major ions (sulfate, nitrate), chemical elements (silicon, zinc)
- Air pollution is generated by both anthropogenic and natural sources
- Source-specific pollution likely varies by source in its associations with adverse health outcomes

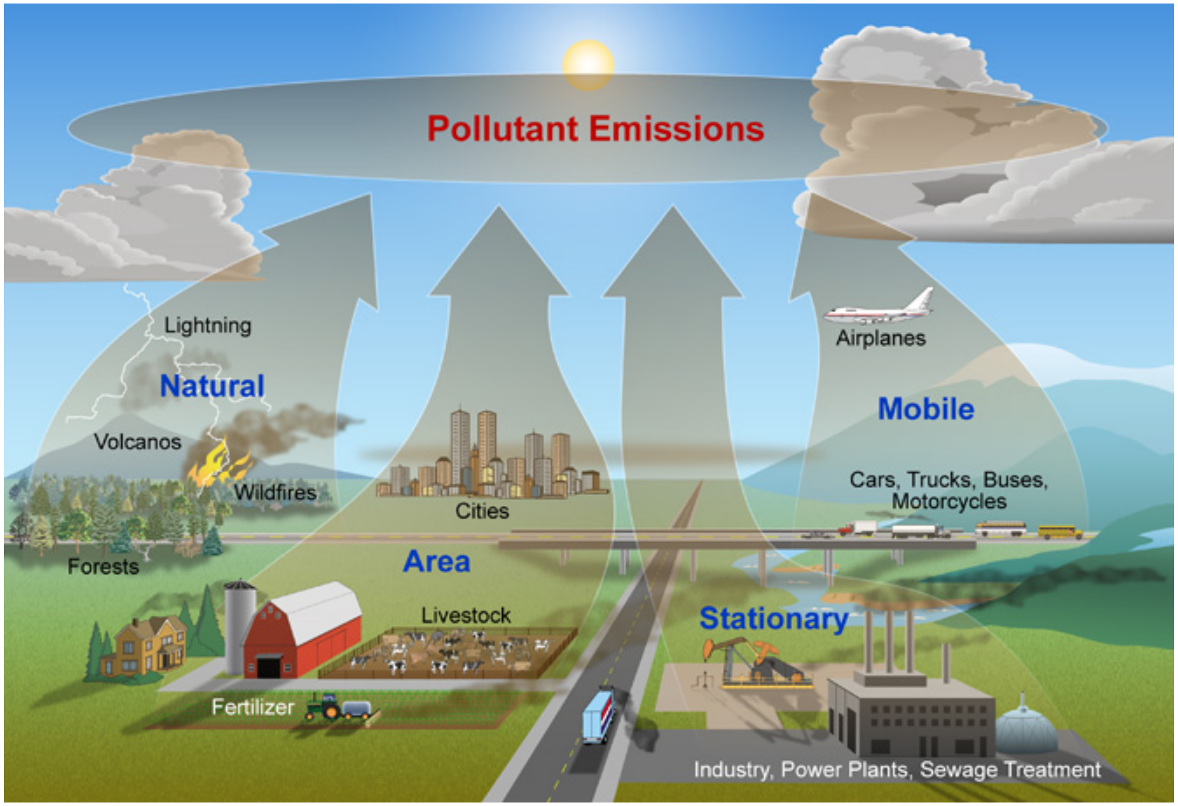


Image: nps.gov

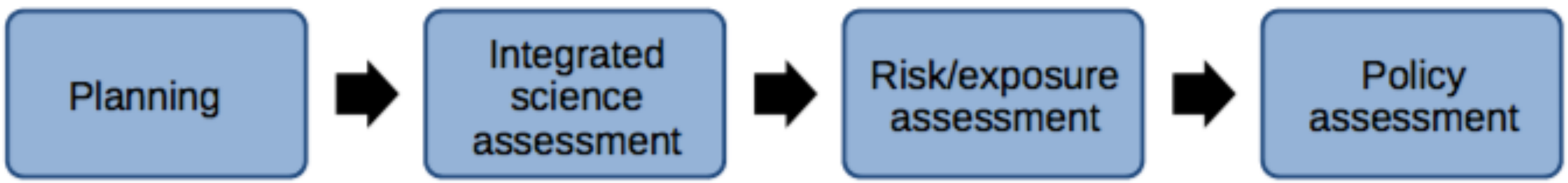
National Ambient Air Quality Standards (NAAQS)

Pollutant [links to historical tables of NAAQS reviews]		Primary/ Secondary	Averaging Time	Level	Form
Carbon Monoxide (CO)		primary	8 hours	9 ppm	Not to be exceeded more than once per year
			1-hour	35 ppm	
Lead (Pb)		primary and secondary	Rolling 3 month period	$0.15\ \mu g/m^3$ (1)	Not to be exceeded
Nitrogen Dioxide (NO_2)		primary	1-hour	100 ppb	98th percentile of 1-hour daily maximum concentrations, averaged over 3 years
		primary and secondary	1 year	53 ppb	Annual Mean
Ozone (O_3)		primary and secondary	8 hours	0.070 ppm (2)	Annual fourth-highest daily maximum 8-hour concentration, averaged over 3 years
Particle Pollution (PM)	$PM_{2.5}$	primary	1 year	$12.0\ \mu g/m^3$	annual mean, averaged over 3 years
		secondary	1 year	$15.0\ \mu g/m^3$	annual mean, averaged over 3 years
		primary and secondary	24 hours	$35\ \mu g/m^3$	98th percentile, averaged over 3 years
	PM_{10}	primary and secondary	24 hours	$150\ \mu g/m^3$	Not to be exceeded more than once per year on average over 3 years

Image: epa.gov

NAAQS review

Separately for each criteria pollutant:



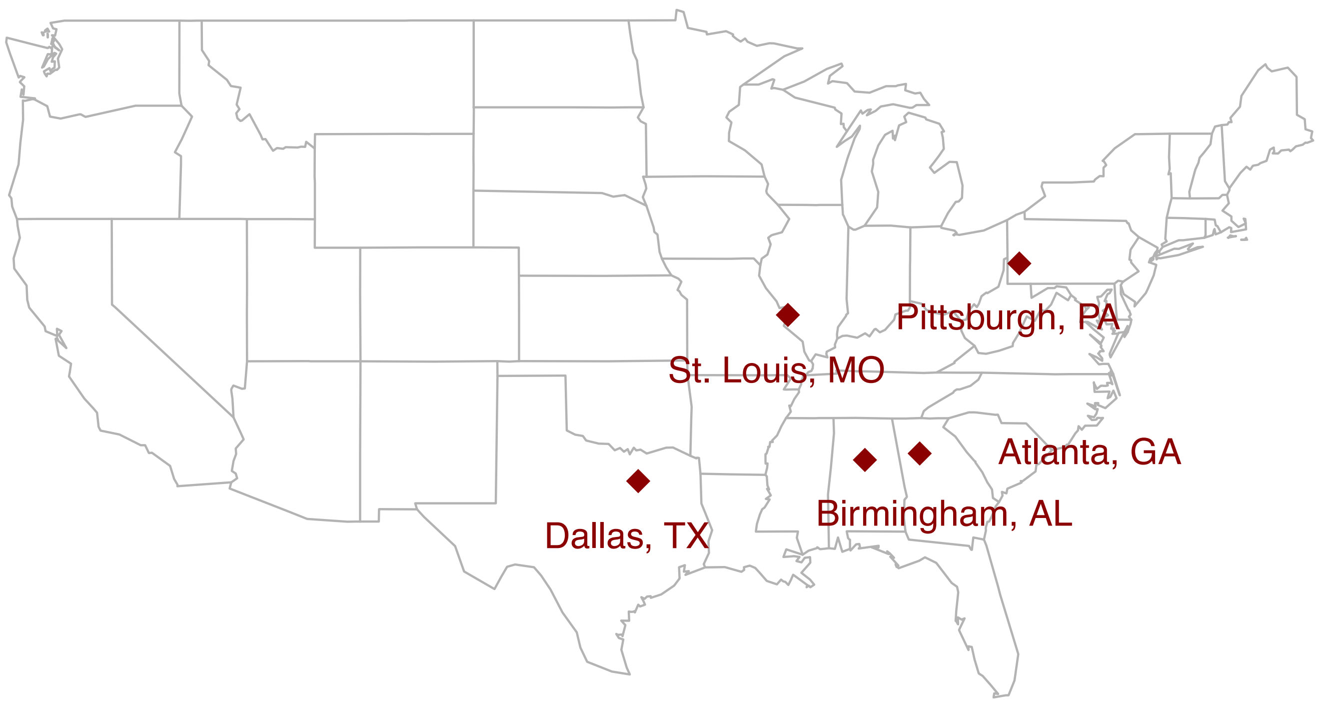
- Mixtures emitted from sources
  - Interpretable
  - Better targets of intervention
- Challenges
  - Sources of pollution are generally estimated and not observed
  - Multicity studies necessary for understanding how pollution impacts health

Multicity Morbidity Study

12 ambient pollutants
Gases: CO, NO_2, NO_x, ozone, SO_2
Particles: PM_{10}, $PM_{2.5}$ and $PM_{2.5}$ constituents EC, OC, NH_4, NO_3, SO_4

→

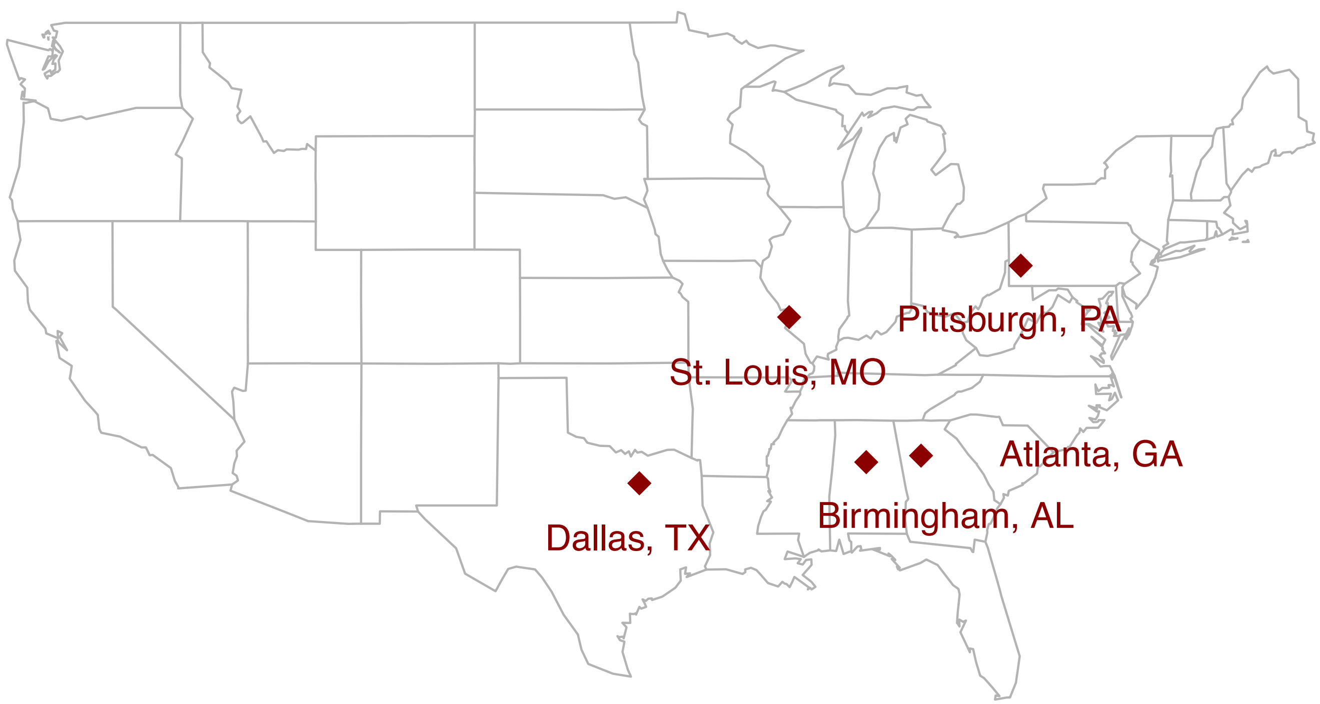
Emergency department (ED) visits for cardiorespiratory diseases



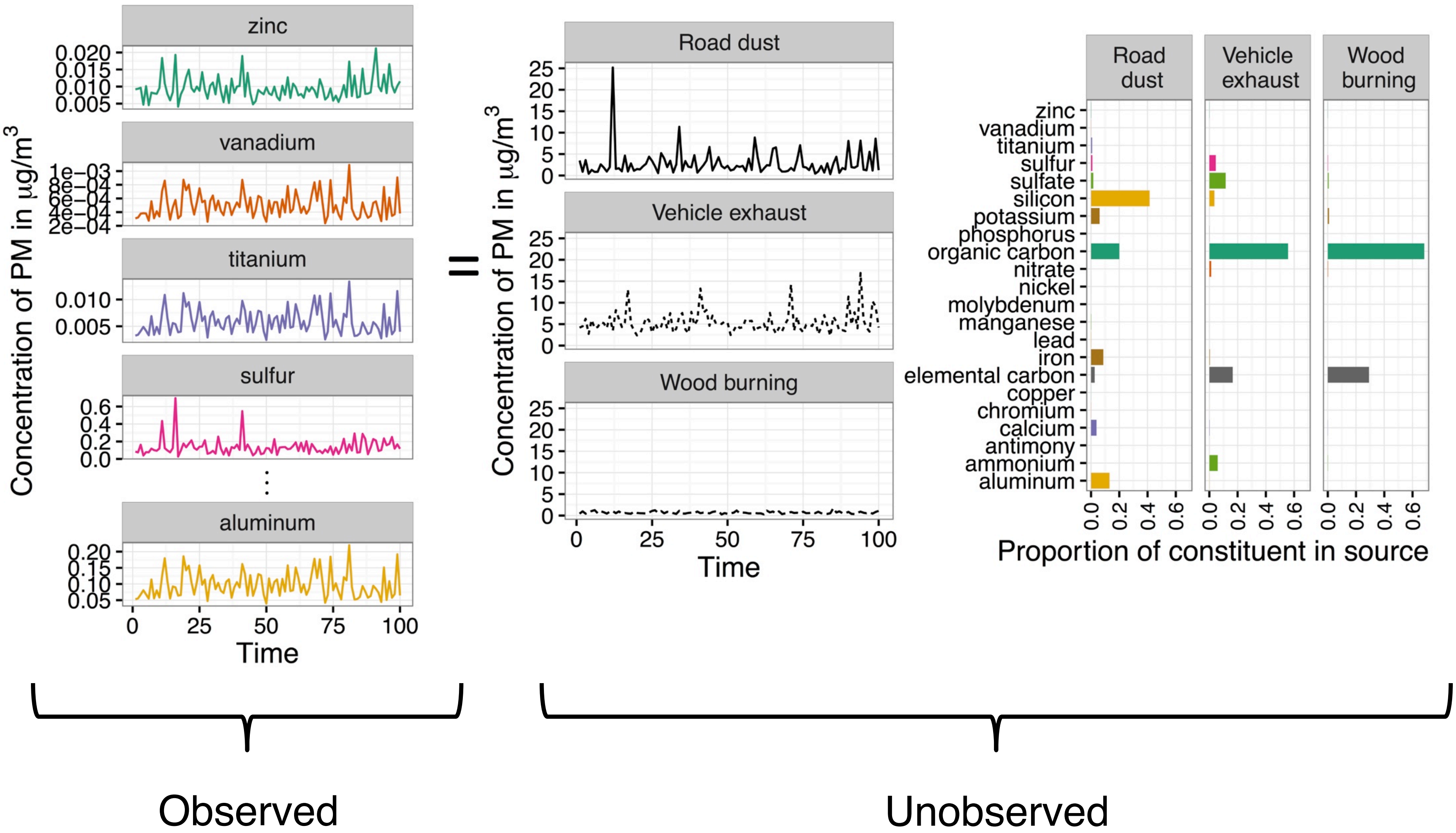
Multicity Morbidity Study

We will compare across 5 US cities:

- Multipollutant factors
- Associations between multipollutant factors and emergency department (ED) visits for cardiorespiratory diseases



Source apportionment models for one city



Adapted from Krall and Strickland (2017) *Current Environmental Health Reports*

Source apportionment models for one city

Estimate unknown source concentrations *F* and source profiles *Λ* from observed data *X*

$$X_{[T \times P]} = F_{[T \times L]} \Lambda_{[L \times P]} + \epsilon_{[T \times P]}$$

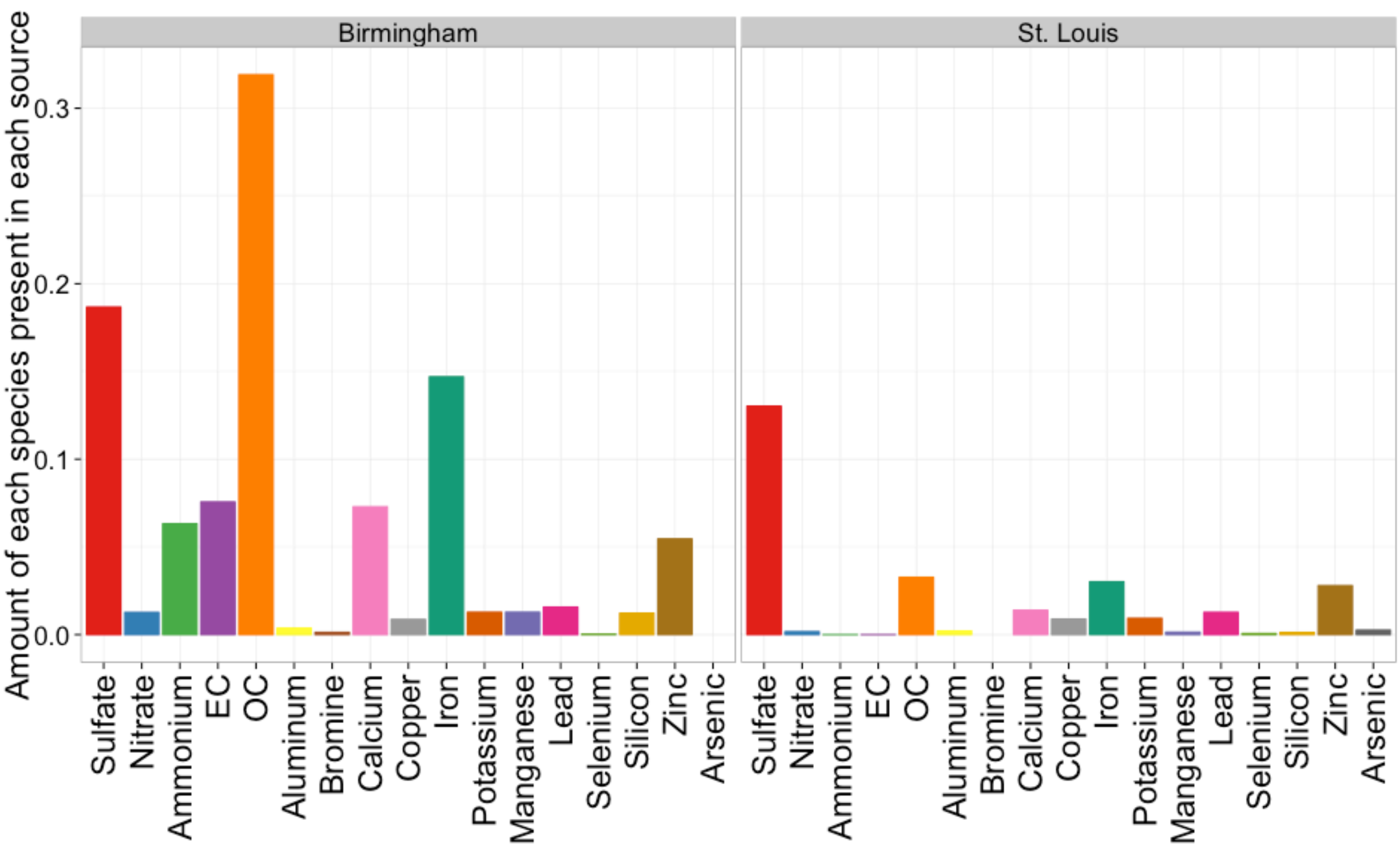
- $x_{t,p}$ Concentration of pollutant *p* on day *t*
- $f_{t,l}$ Concentration of source *l* on day *t*, ≥ 0
- $\lambda_{l,p}$ Amount pollutant *p* contributes to source *l*, ≥ 0
  - Generally we assume that $\Sigma_p \lambda_{l,p} = 1$
- $\varepsilon_{t,p}$ Measurement error or unexplained pollution

$\mathbf{f}_l$ used in studies of the short-term associations between pollution from source *l* and acute health outcomes for a single community.

Challenges for multicity studies

- Source-specific pollution is estimated separately for each city
- Sources-specific pollution varies in chemical composition between cities

$PM_{2.5}$ from metals source

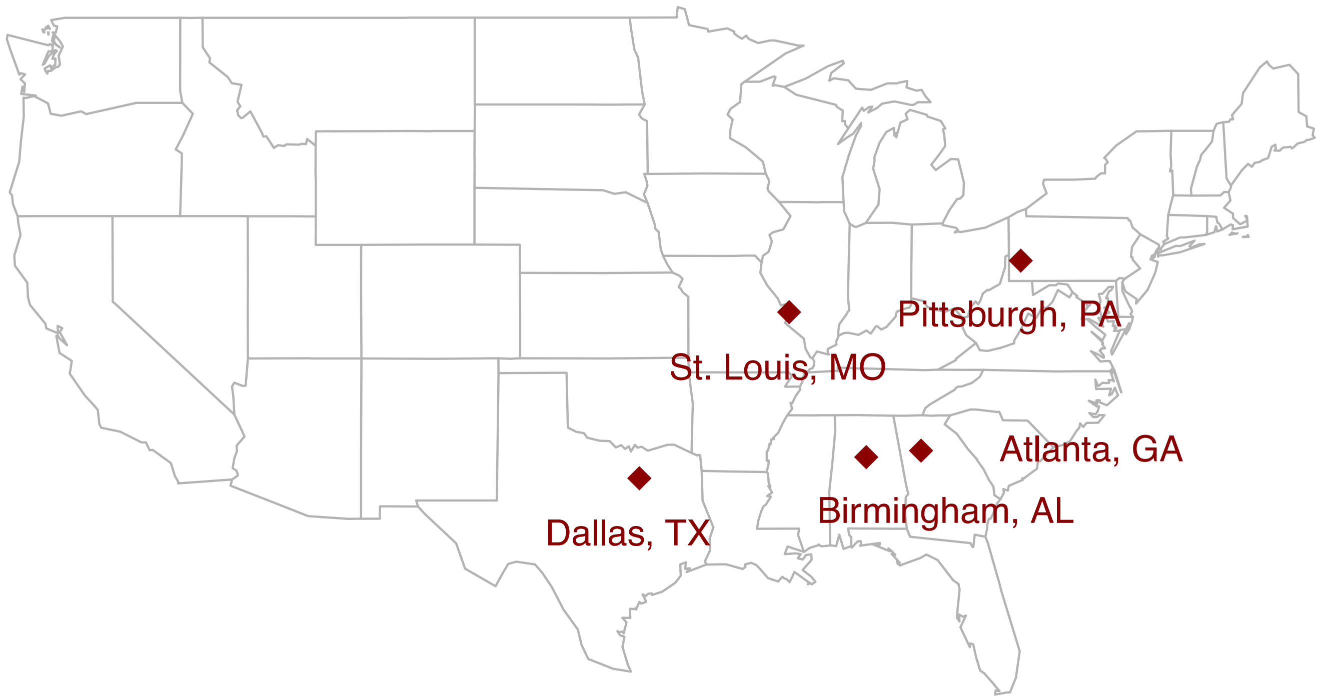


Challenges for multicity epidemiologic studies

Between-city heterogeneity in estimated health effects may be driven by

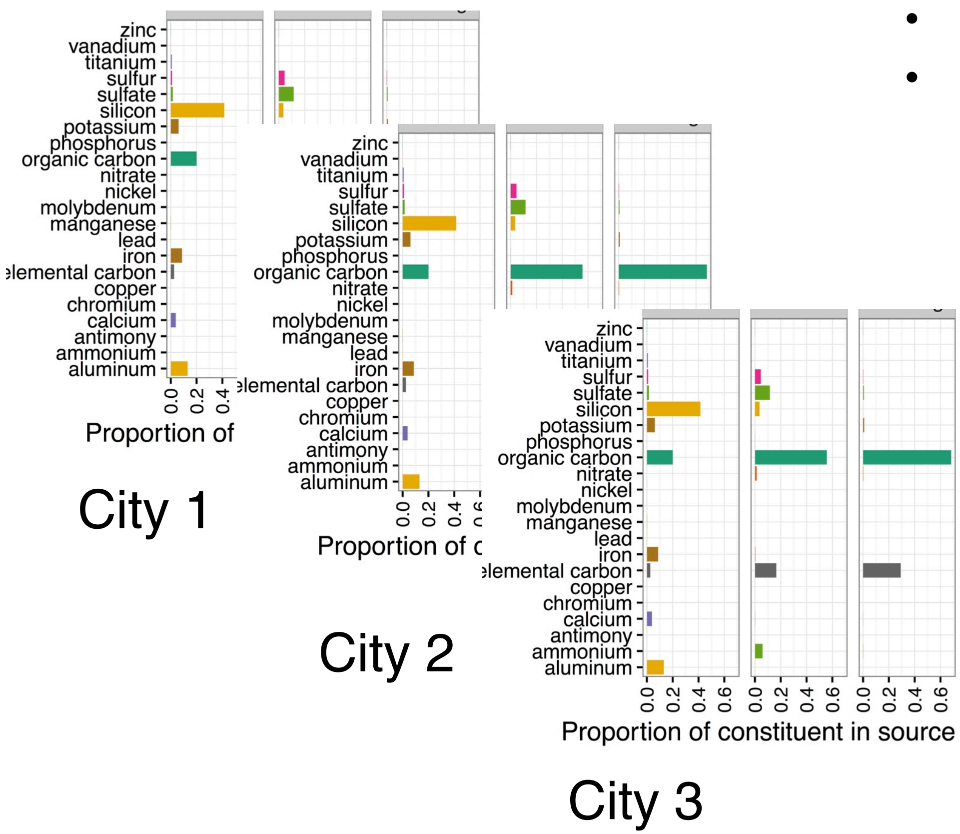
- Differences in population or exposure characteristics
- Differences in pollution composition

What population characteristics drive between-city differences?

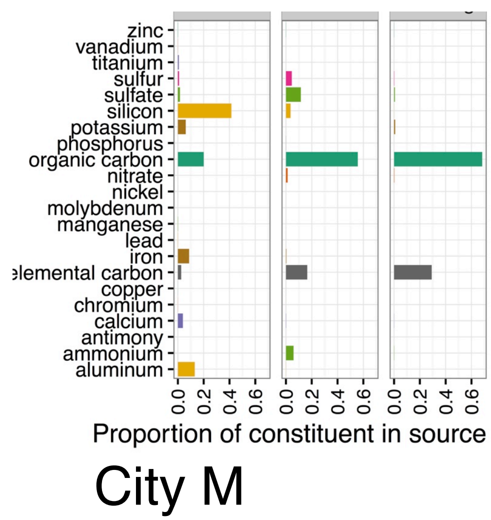


Source apportionment in multicity studies

- Differences across cities in:
  - Numbers of sources
  - Chemical composition of sources

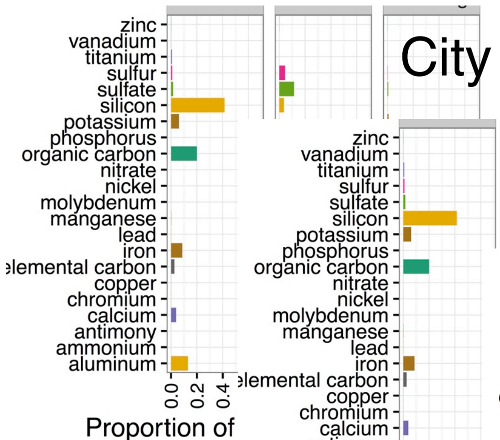


…



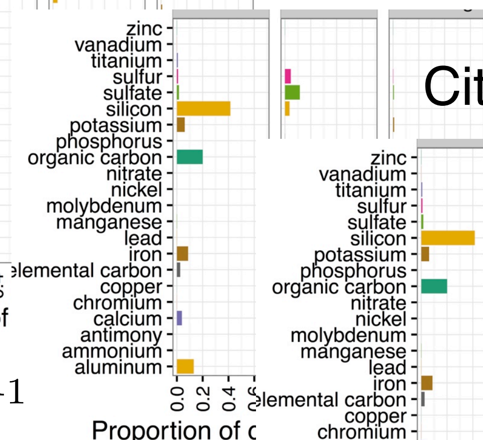
Source apportionment in multicity studies

City 1



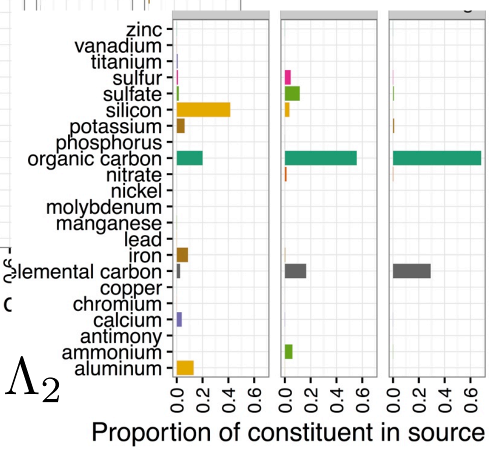
$X_1 \approx F_1 \Lambda_1$

City 2



$X_2 \approx F_2 \Lambda_2$

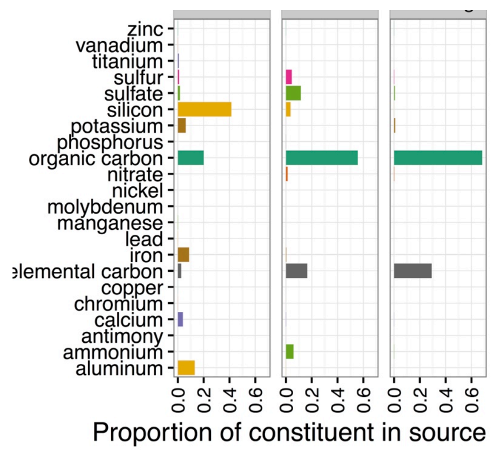
City 3



$X_3 \approx F_3 \Lambda_3$

…

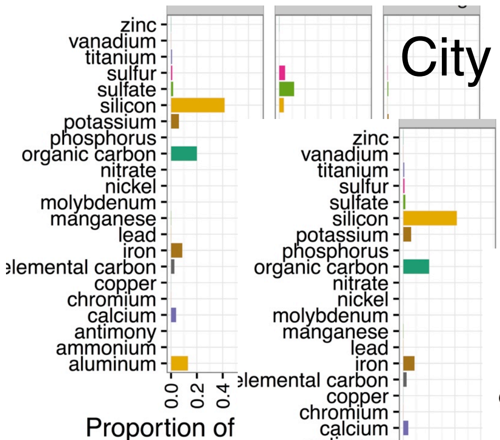
City M



$X_M \approx F_M \Lambda_M$

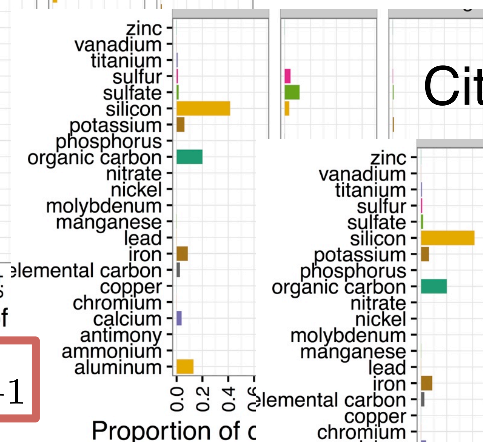
Source apportionment in multicity studies

City 1



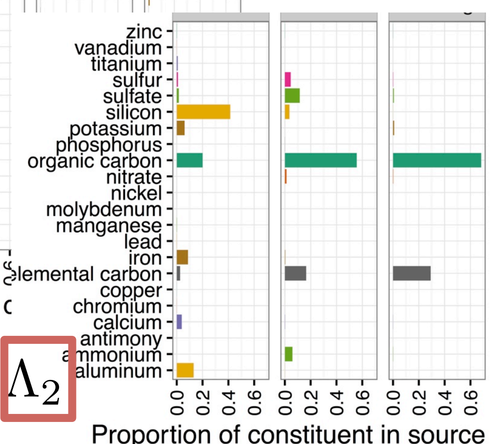
$$X_1 \approx F_1 \boxed{\Lambda_1}$$

City 2



$$X_2 \approx F_2 \boxed{\Lambda_2}$$

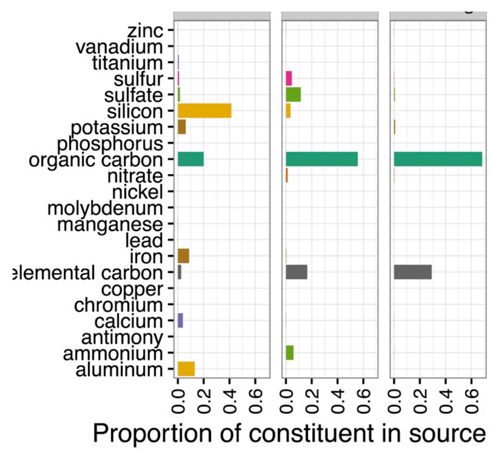
City 3



$$X_3 \approx F_3 \boxed{\Lambda_3}$$

...

City M



$$X_M \approx F_M \boxed{\Lambda_M}$$

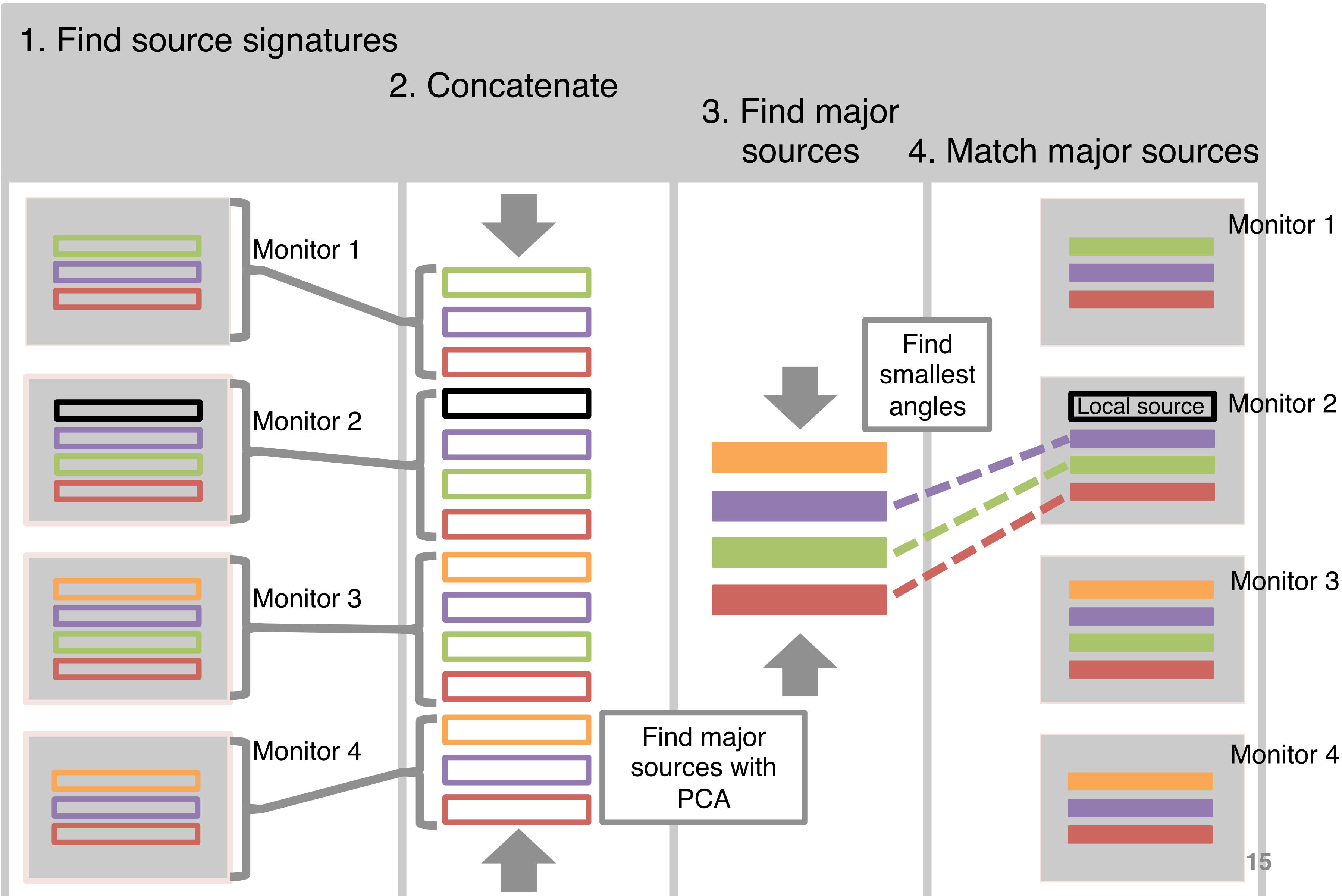
Methods: SHARE combines sources across cities

SHARE is a ***population value decomposition*** approach for combining source estimates across cities

- SHARE determines pollution factors that are shared across a region.
- For city *i* (*i = 1,…,5*)
  - Single-community factor analysis: $X_i \approx F_i \Lambda_i$
  - Proposed population value decomposition: $X_i \approx F_i^* \Lambda$
- The SHARE approach leverages:
  - City-specific source concentrations (exposure): F_i^*
  - Population level latent factors (major factors): Λ

Krall, Hackstadt, & Peng (2017) *Statistics in Medicine*

Methods: SHARE combines sources across cities



Methods: Time series health models

Estimate associations between each multipollutant factor and diagnosis separately for each city using overdispersed Poisson regression models:

$$log(\mu_{tjc}) = \beta_0 + \beta_{cjl} F_{t'cl} + \text{confounders}$$

μ_{tjc} E(Y_{tjc})
Y_{tjc} Number of ED visits for day *t*, city *c*, and diagnosis *j*.
$F_{t'lc}$ Concentration for day *t'*, city *c*, and multipollutant factor *l*.
β_{cjl} Log relative risk for city *c*, diagnosis *j*, and multipollutant factor *l.*

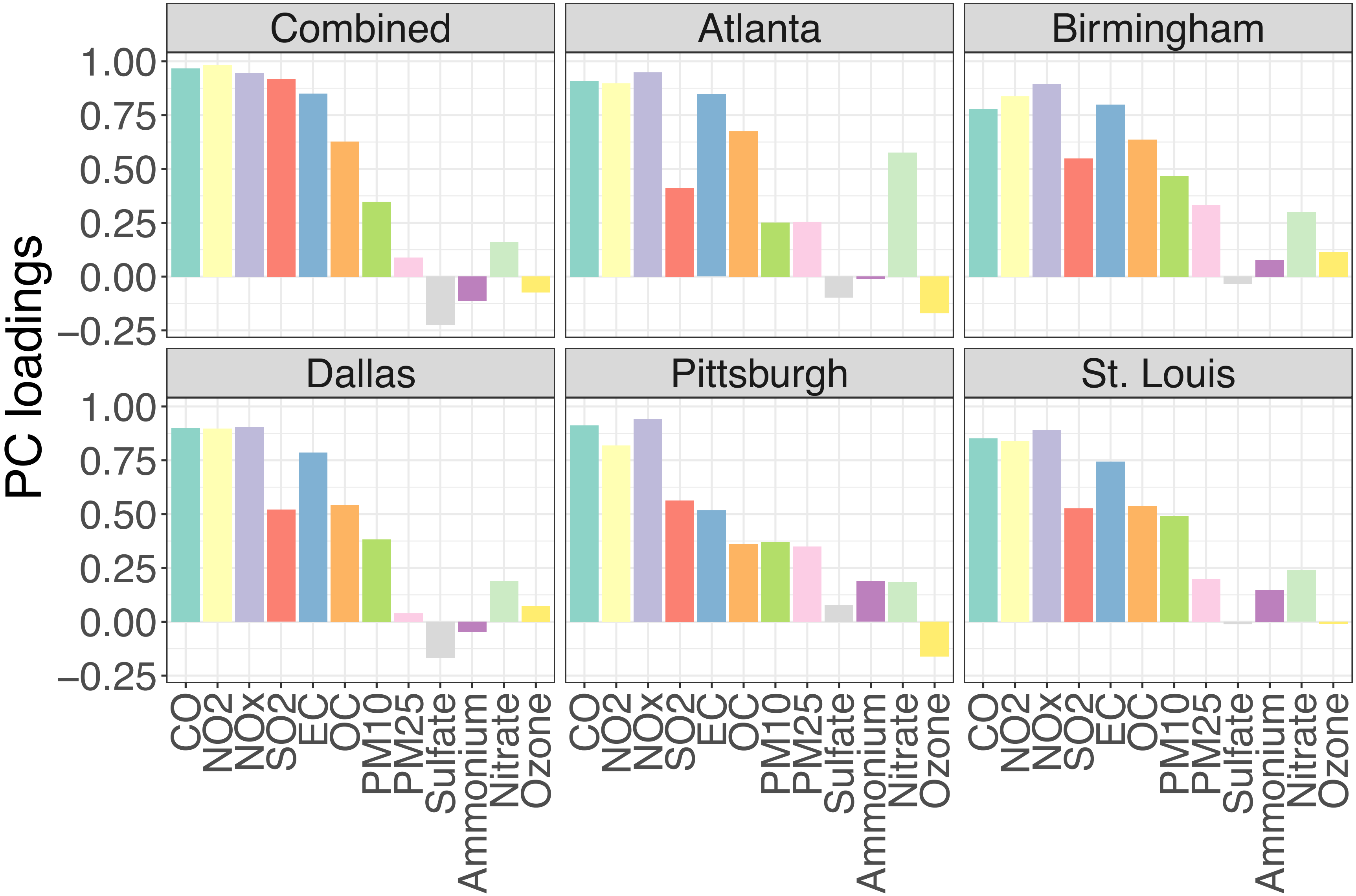
Confounder control:

- Holidays
- Day of week
- Season
- Cubic terms for maximum and mean temperature
- Cubic terms for dew point temperature
- Temporal trends

Lag of exposure:
Cardiovascular: same day (lag 0) exposure
Respiratory: mean 0-7 days exposure

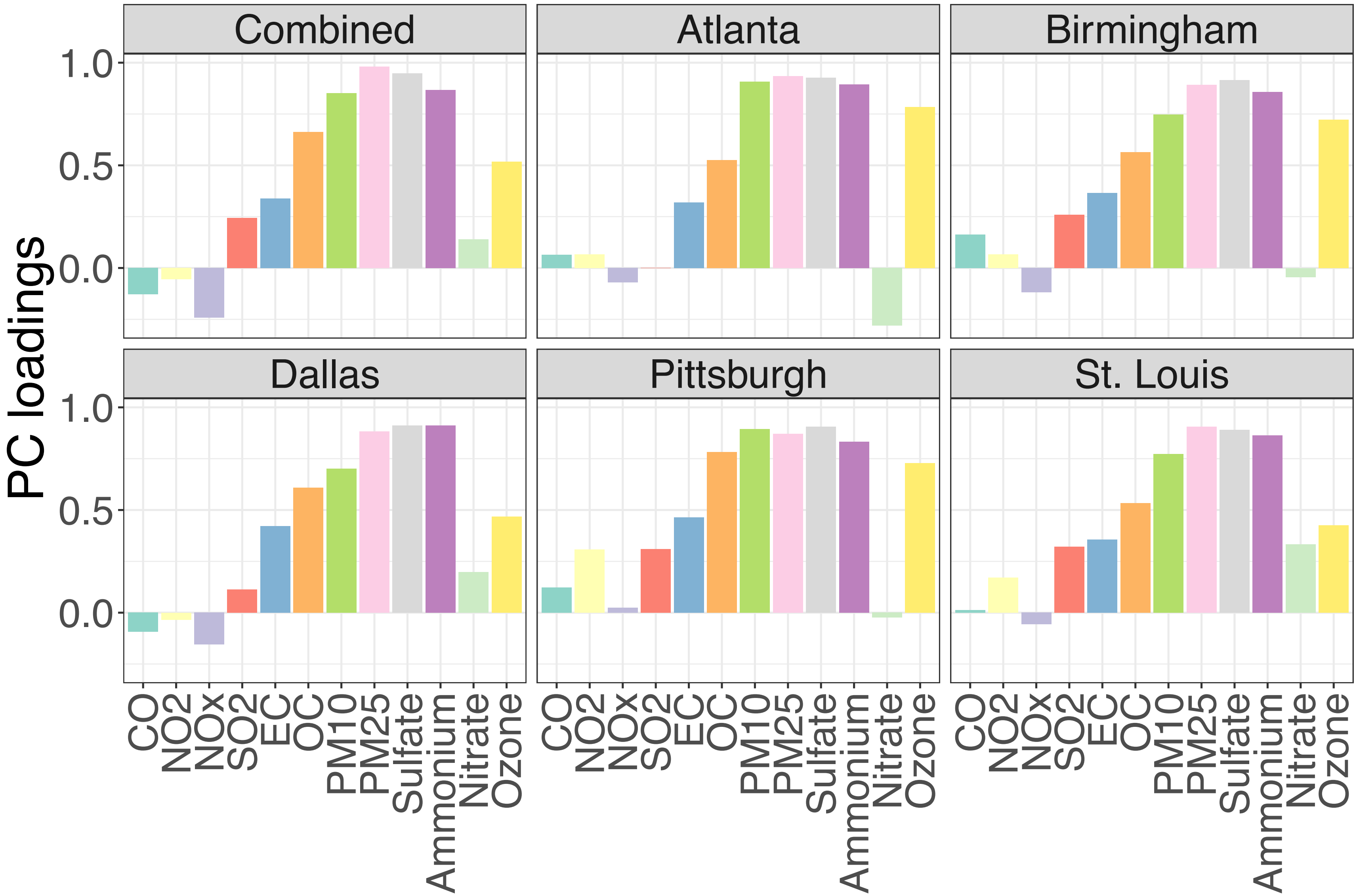
Results: Multicity morbidity study

Factor 1: Primary pollution



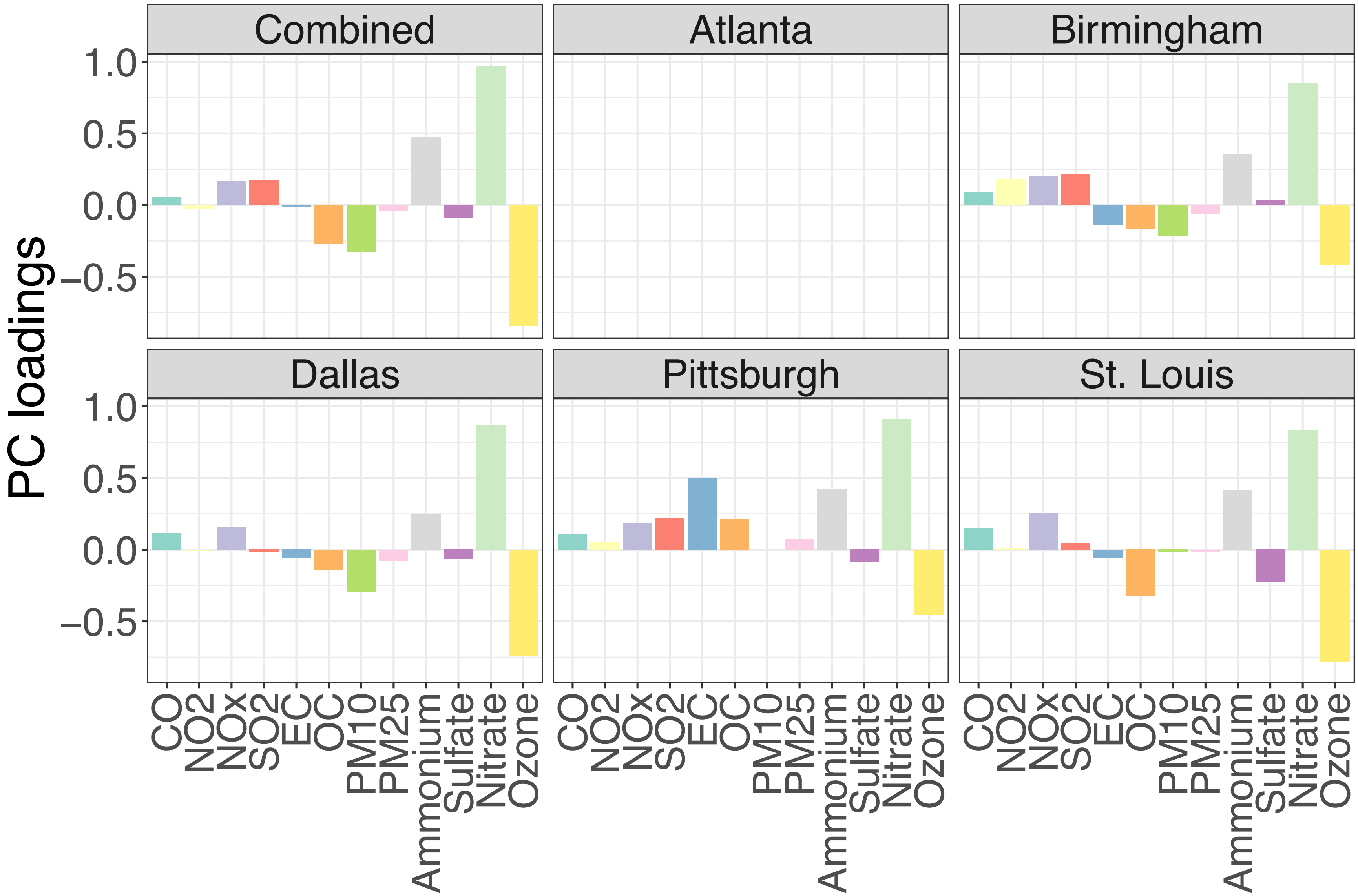
Results: Multicity morbidity study

Factor 2: Secondary pollution

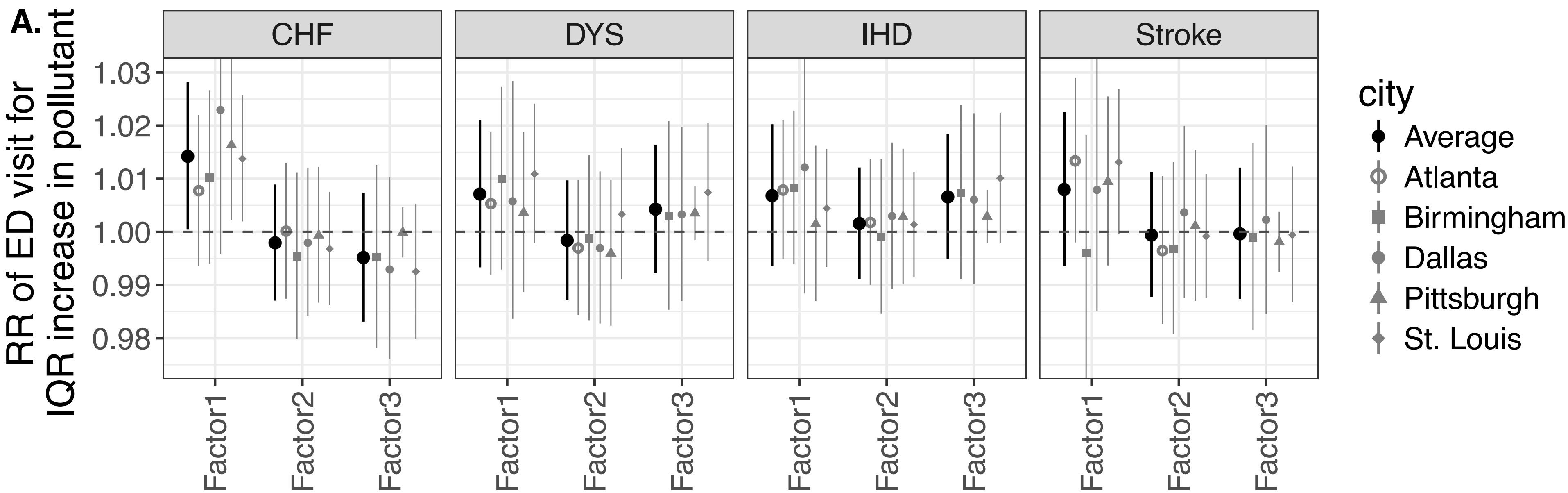


Results: Multicity morbidity study

Factor 3: Secondary nitrate



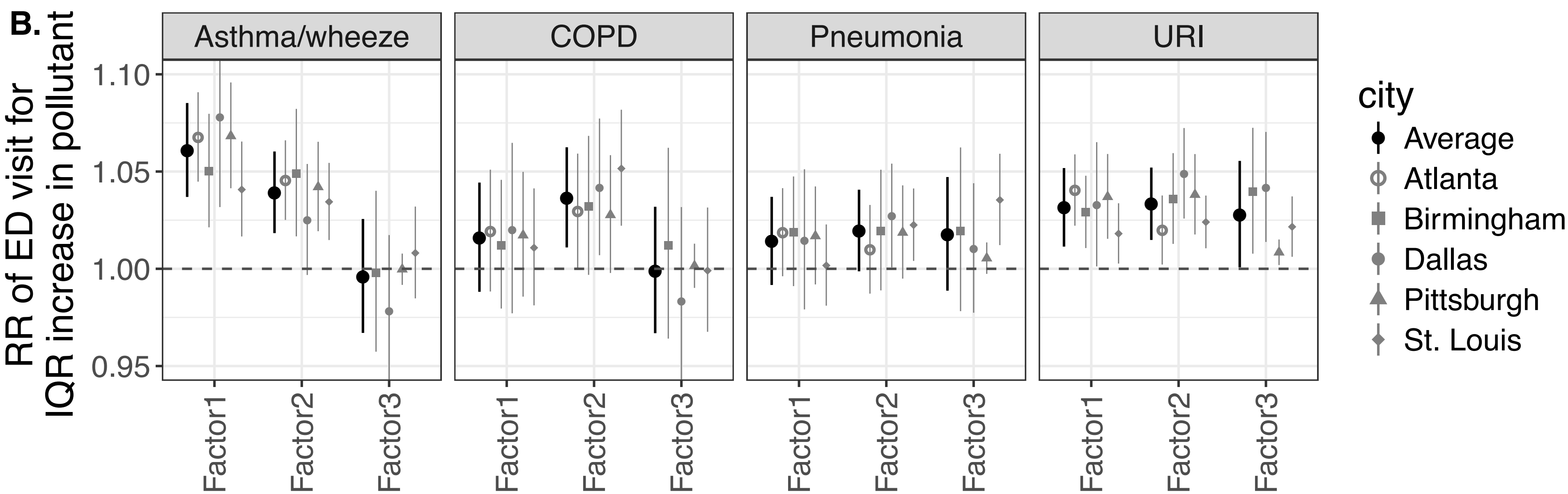
Results: Multicity morbidity study



- CHF: congestive heart failure
- DYS: cardiac dysrhythmia
- IHD: ischemic heart disease
- Stroke

20

Results: Multicity morbidity study



- Asthma and/or wheeze
- COPD: chronic obstructive pulmonary disease
- Pneumonia
- URI: upper respiratory infection

21

Conclusions and future work

Conclusions

1. Primary pollution might be more associated with cardiovascular diseases, including congestive heart failure and stroke.
2. Both primary and secondary pollution were associated with respiratory diseases, including asthma/wheeze and upper respiratory infection.
3. To better identify sources, we need more measures of chemical elements (e.g. zinc).
4. SHARE approach can be used to facilitate multicity studies of source-specific pollution and multipollutant factors.

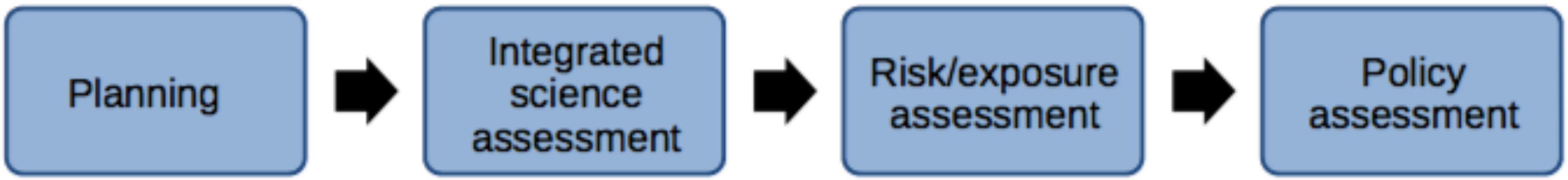
Future work

1. Incorporate more chemical elements into multipollutant factor estimation.
2. Determine threshold for similarity in Λ_i between cities.

Air pollution policies

Both primary and secondary pollution were associated with respiratory ED visits for asthma/wheeze and upper respiratory disease

National ambient air quality standards



Policies aimed at reducing primary pollution have focused on:

- Coal-fired power plants
- Vehicle exhaust emission standards
- Solid waste incinerators
- Others

Thanks!

Jenna Krall

jkrall@gmu.edu

This research is funded by
U.S. EPA - Science To Achieve Results (STAR) Program
Grant # R834799

Acknowledgements

Emory University
Stefanie Ebelt Sarnat
Andrea Winquist
Paige E. Tolbert
Lance A. Waller
Howard Chang

Georgia Tech
James A. Mulholland
Armistead G. Russell

University of Pittsburgh
Evelyn Talbott
Judith Rager

This presentation was made possible by grants from NIEHS (T32-ES012160) and US EPA (R834799). This publication's contents are solely the responsibility of the grantee and do not necessarily represent the official views of the US EPA. Further, US EPA does not endorse the purchase of any commercial products or services mentioned in the publication.

Methods: SHARE combines sources across cities

Find population-level matrix Λ such that $\Lambda_i \approx \Lambda_i \Lambda^T \Lambda$

1. Apply PCA to $\tilde{\Lambda} = [\Lambda_1^T, \Lambda_2^T, ..., \Lambda_M^T]^T$ for M ambient monitors.
2. Then $\tilde{\Lambda} = \tilde{\Lambda} W W^T$, where
   - W^T is the matrix of principal component loadings of $\tilde{\Lambda}^T \tilde{\Lambda}$
   - $W W^T$ is the identity matrix
3. Let Λ^T be the matrix of the first L principal component loadings
4. Since $\tilde{\Lambda} \approx \tilde{\Lambda} \Lambda^T \Lambda$, then $\Lambda_i \approx \Lambda_i \Lambda^T \Lambda$.

Using Λ_i and Λ, we can rewrite

$$X_i \approx F_i \Lambda_i \approx (F_i \Lambda_i \Lambda^T) \Lambda = \tilde{F}_i \Lambda$$