Epidemiology of multiple PM components

HEI Annual Conference

Helen H. Suh
Department of Civil & Environmental Engineering
Tufts University

April 29, 2020

Potential Scenarios

PM$_{2.5}$:

- Differential toxicity
- Regional, seasonal differences
- Inconsistent component and source findings

Individual components

Sources

Non-, Low toxicity component

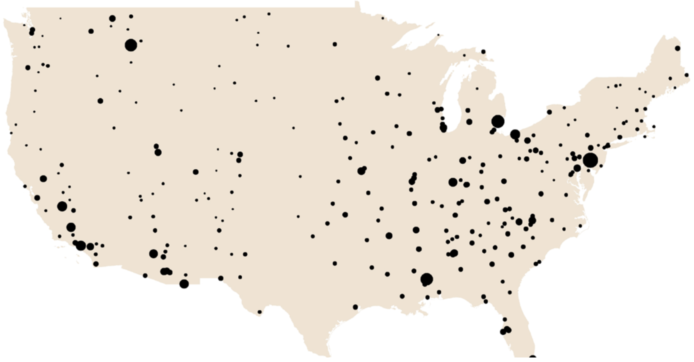
Mixtures

Current

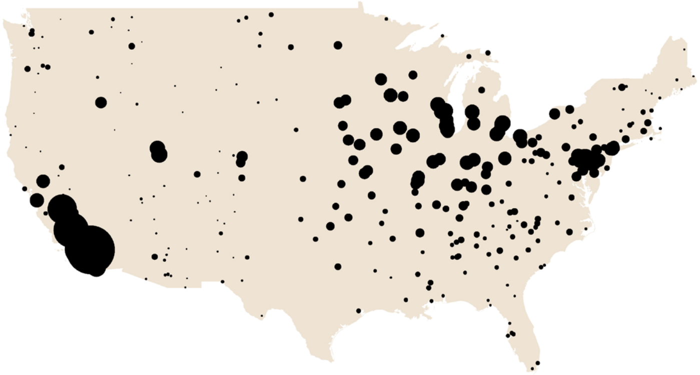
Future?

$PM_{2.5}$ Source-Related Concentrations

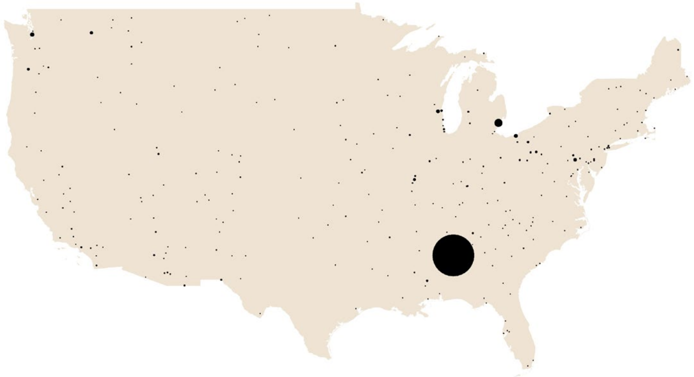
Biomass Combustion



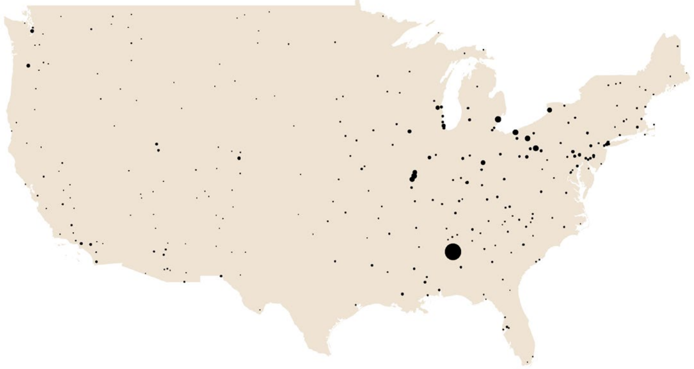
Traffic



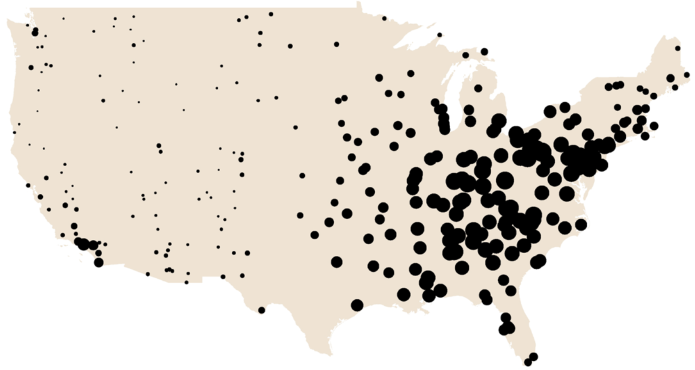
Steel Industry



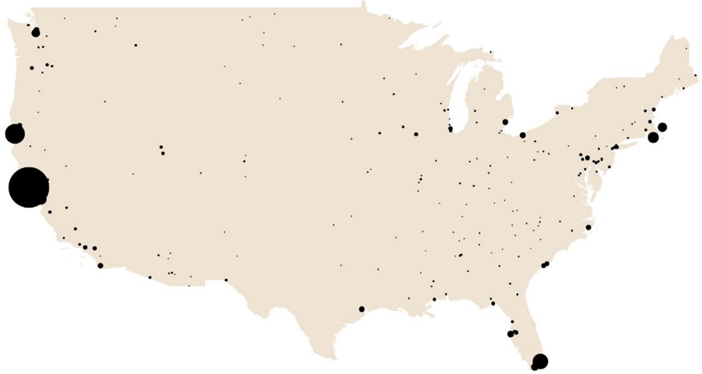
Metals



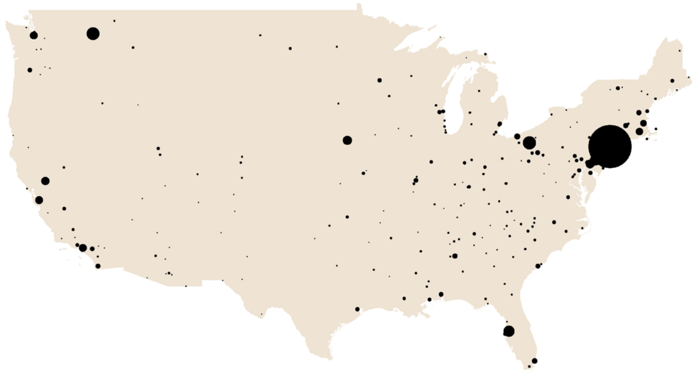
Coal Combustion



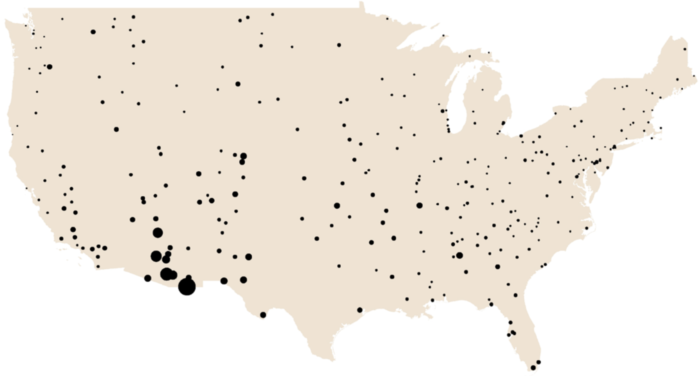
Salt



Residual Oil Combustion

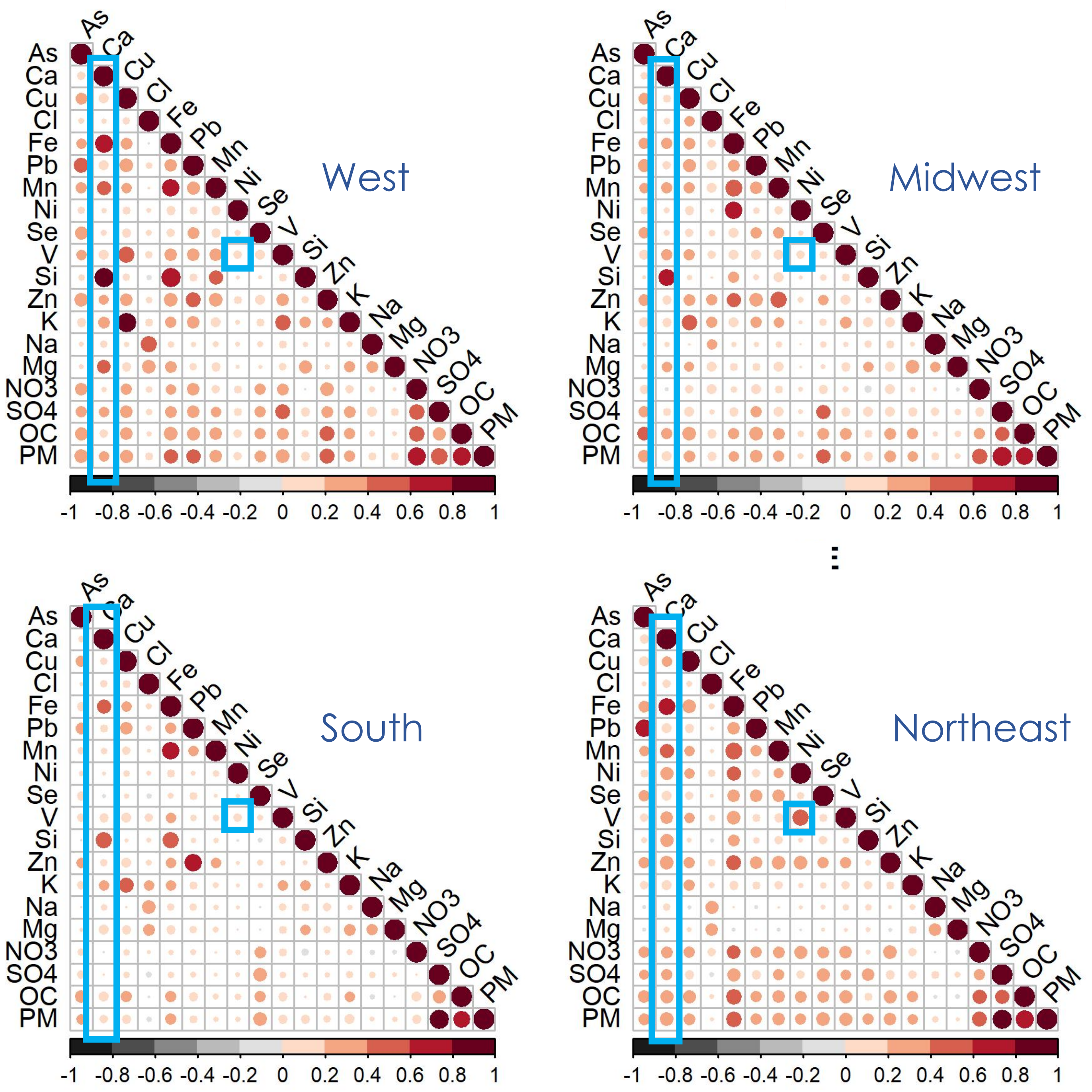


Soil



Each circle radius represents the soil and metal concentrations times 10, traffic times 5, biomass times 25, salt times 100, steel, residual oil times 50; estimated using PMF for 2000-2008.

Correlations among $PM_{2.5}$ components*: by region of country



*Correlations amongst 24-h values

Potential Scenarios

$PM_{2.5}$:

- Differential toxicity
- Regional, seasonal differences
- Inconsistent component and source findings

Individual components

Sources

Non-, Low toxicity component

Mixtures

Current

Future?

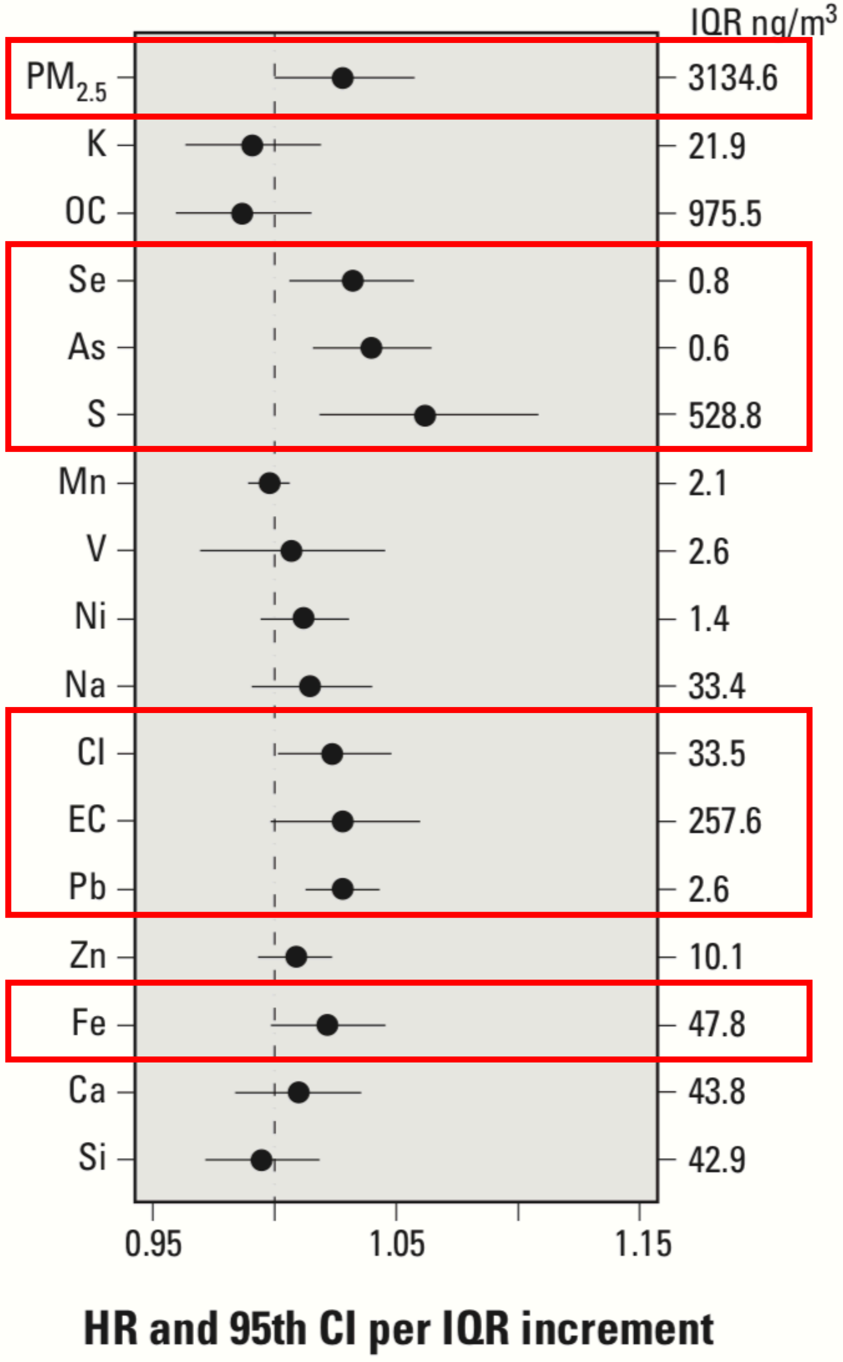
Association of IHD mortality and $PM_{2.5}$ elements[1]

Cohort: ACS CPS-II (~450,000 adults)

Outcome: IHD mortality

Exposures: Multi-year MA $PM_{2.5}$, components for 100 MSAs

Analysis: Random effects Cox proportional hazards, adjusting for individual and contextual variables



[1]Thurston et al., 2016, *EHP*

Association of CVD mortality and $PM_{2.5}$ elements[1]

Cohort: WHI-OS (~73k postmenopausal women)

Outcome: CVD mortality

Exposures: 1-year exposures using spatial model

Analysis: Cox proportional hazard models, adjusting for individual level covariates

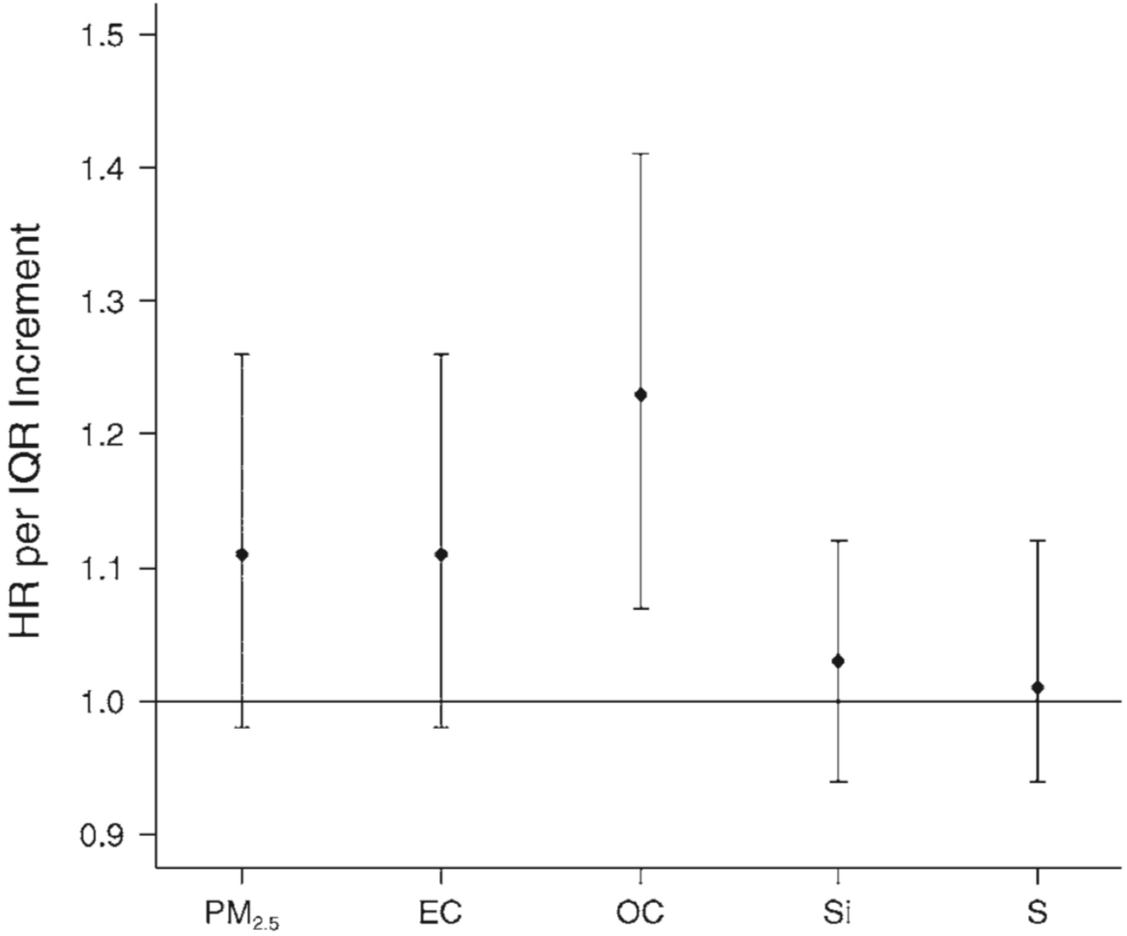


Figure 41. Estimated HR per IQR increment in $PM_{2.5}$ and $PM_{2.5}$ components for CVD death (95% CI). See Table 52 for details.

Association of Cardiopulmonary, IHD mortality and $PM_{2.5}$ components[1]

Cohort: CA Teachers Study (~43k women)

Outcome: All-, CVD mortality

Exposures: Long-term exposure, 30km of monitoring site

Analysis: Cox proportional hazard, adjusting for individual, area covariates

	Cardiopulmonary (ICD-10 codes, I00–I99, J00–J98) (*n* = 1,357)		Ischemic heart disease (ICD-10 codes, I20–I25) (*n* = 460)	
Pollutant	**HR (95% CI)**	***p*-Value**	**HR (95% CI)**	***p*-Value**
$PM_{2.5}$	1.11 (1.03–1.21)	0.01	1.31 (1.14–1.50)	< 0.01
EC	1.07 (0.94–1.22)	0.28	1.46 (1.17–1.83)	< 0.01
OC	1.04 (0.98–1.11)	0.19	1.13 (1.01–1.25)	0.03
Sulfate	1.14 (1.01–1.29)	0.03	1.48 (1.20–1.82)	< 0.01
Nitrate	1.11 (1.03–1.19)	0.01	1.27 (1.12–1.43)	< 0.01
Iron	1.05 (0.93–1.19)	0.40	1.39 (1.13–1.72)	< 0.01
Potassium	1.06 (0.97–1.17)	0.22	1.27 (1.07–1.49)	< 0.01
Silicon	1.05 (1.00–1.10)	0.04	1.11 (1.02–1.20)	0.01
Zinc	1.09 (0.98–1.20)	0.10	1.33 (1.12–1.58)	< 0.01

[1]Ostro et al. (2011), EHP

Single pollutant model: concerns

- Risk estimate reflects effect of pollutant AND correlated co-pollutants
- Many $PM_{2.5}$ constituent concentrations are low, and have higher associated measurement error → may result in unstable effect estimates
- Difficult to separate health effects of constituents that comprise large fraction of $PM_{2.5}$, e.g., sulfur, from $PM_{2.5}$

➢ Interpretation of model coefficients is complicated

Potential Scenarios

$PM_{2.5}$:

- Differential toxicity
- Regional, seasonal differences
- Inconsistent component and source findings

Individual components

Sources

Non-, Low toxicity component

Mixtures

Current

Future?

Source Characterization

- **Source tracers:** individual pollutants reflect exposures for specific sources
- **Spatial models:** estimates based on location-specific parameters, can reflect exposures to one or more source types
- **Back trajectory analysis**: days or periods clustered based on the origin and route of the air parcels, representative of different pollutant-profiles
- **Source apportionment**: groups pollutants by how they co-vary from day-to-day to represent common sources

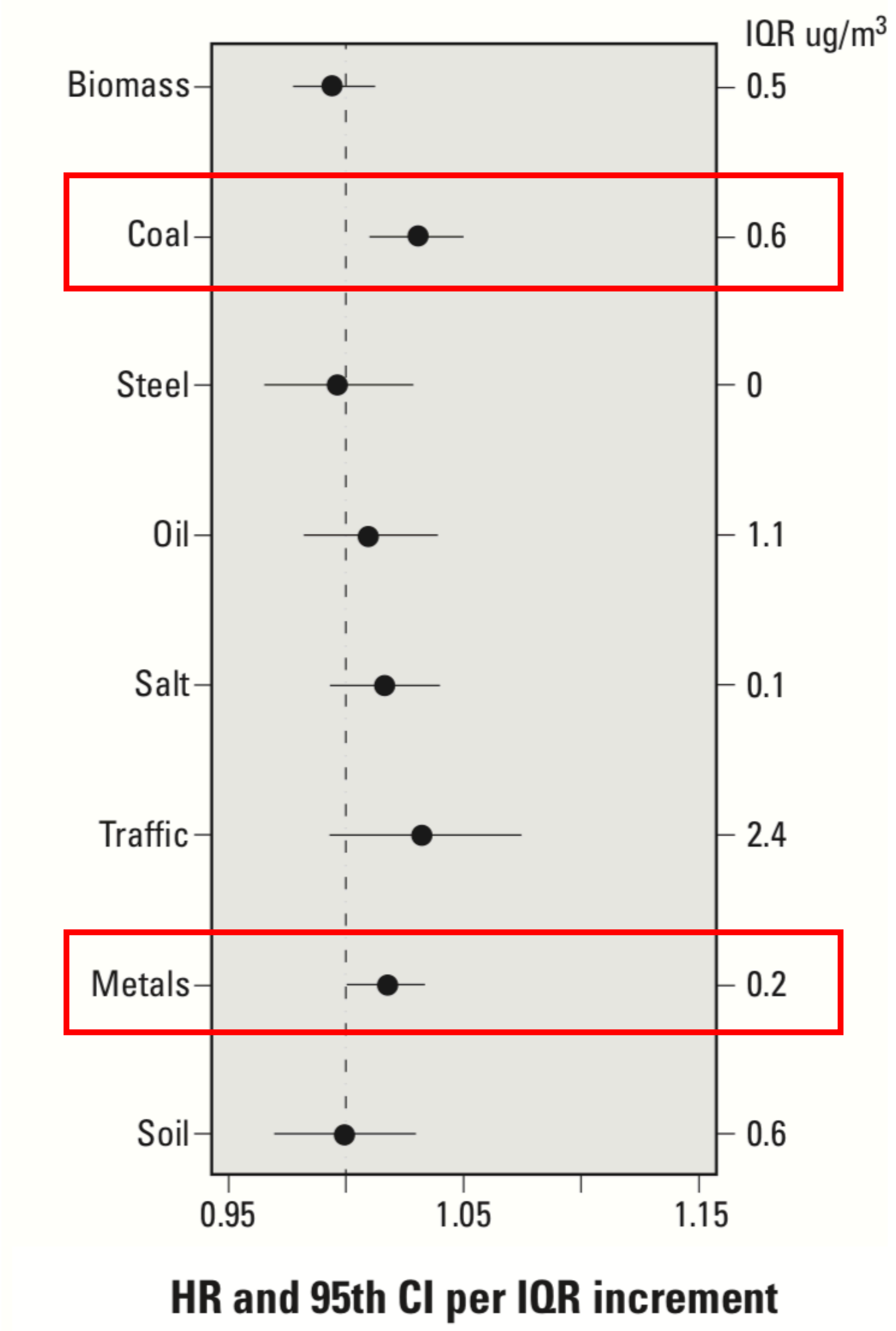
$PM_{2.5}$ sources and IHD mortality[1]

Cohort: ACS CPS-II (~450,000 adults)

Outcome: IHD mortality

Exposures: SA based on multi-year MA $PM_{2.5}$, components for 100 MSAs

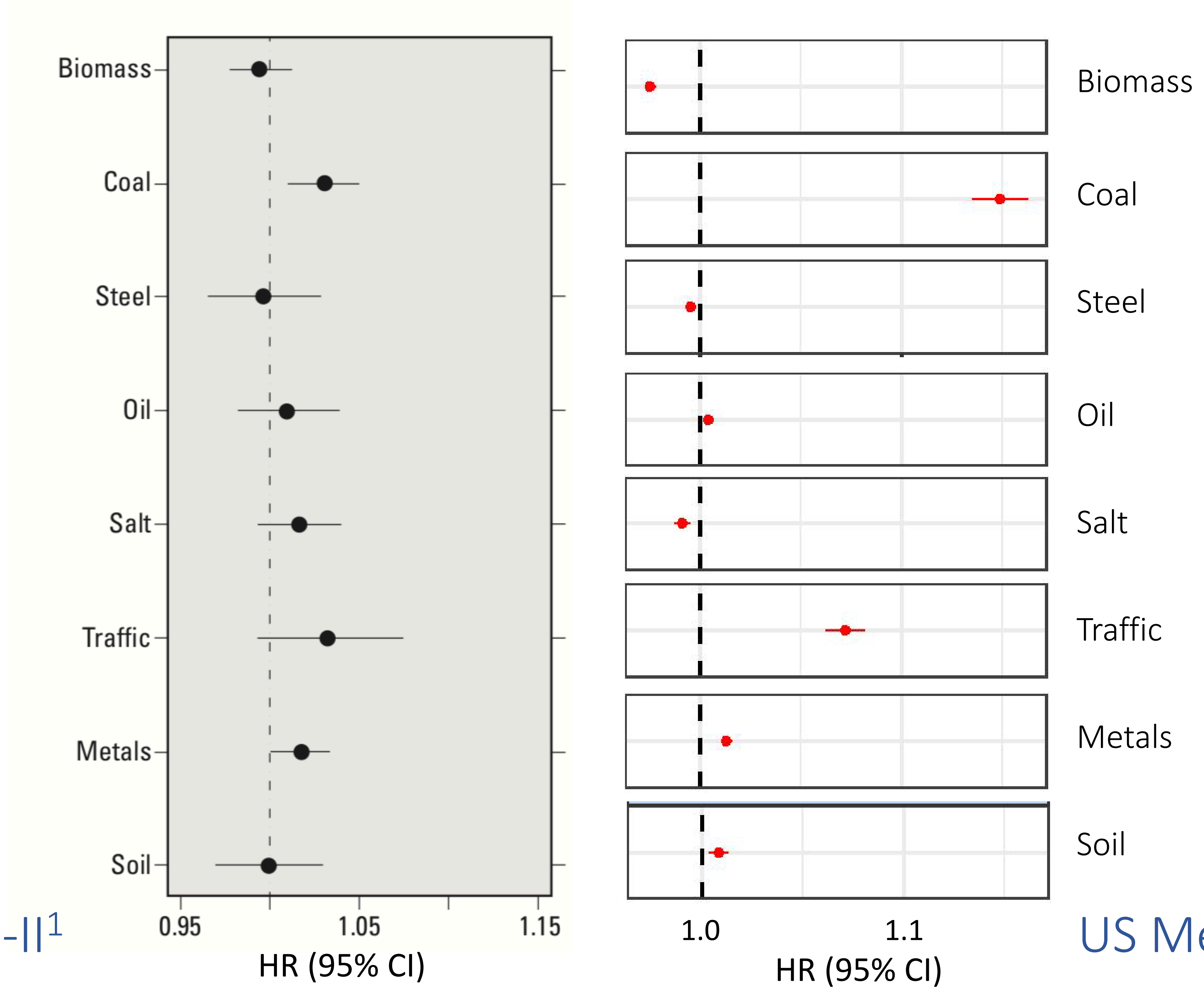
Analysis: Random effects Cox proportional hazards, adjusting for individual and contextual variables



[1]Thurston et al., 2016, *EHP*

$PM_{2.5}$ sources and IHD mortality

[1]Thurston et al., 2016, *EHP*
[2]Kazemiparkouhi et al. 2020, *in prep*



ACS CPS-II[1]

US Medicare[2]

Potential Scenarios

$PM_{2.5}$:
- Differential toxicity
- Regional, seasonal differences
- Inconsistent component and source findings

Individual components

Sources

Non-, Low toxicity component

Mixtures

Current

Future?

Possible approaches

Recent approaches aimed at explicitly assessing multiple components simultaneously and their complex relationships

- Residual-based method (Mostofsky et al., 2012; Wang et al. 2020)
- Bayesian kernel machine regression (BKMR) (Bobb et al., 2015; 2018)
- Bayesian hierarchical models (Chung et al., 2015)
- Lasso regression (Peng et al., 2017)

➢ For methods to be useful, must be (1) validated, (2) easily and efficiently applied to epidemiological datasets, and (3) produce readily interpretable results

Residual Method

(Mostofsky et al., 2012; Wang et al., 2020)

Goal: Assess whether the impact of $PM_{2.5}$ on health endpoint is reduced when the influence of a specific component(s) is removed

Approach: Use 2-stage approach

- *Stage I:* $PM_{2.5} = \alpha + \beta * Component(s)$ by location
- *Stage II:* Residual of Stage 1 as exposure measure

Mortality and Total and Non-Traffic $PM_{2.5}$

Cohort: US Medicare beneficiaries (~53 million older adults)

Outcome: Cause-specific mortality

Exposures: $PM_{2.5}$, NO_2 conterminous US

Analysis: Residual 2-stage model, hybrid machine learning-Cox proportional hazards, adjusted for age, gender, race, and area SES

Non-traffic $PM_{2.5}$: Residual of $PM_{2.5}$~NO_2

Causes of Death	$PM_{2.5}$	Non-Traffic $PM_{2.5}$
All Causes	1.050 (1.044,1.056)	1.015 (1.008,1.022)
Non-accidental	1.051 (1.045,1.057)	1.014 (1.007,1.021)
Accidental	0.998 (0.960,1.037)	1.018 (0.972,1.066)
All Cardiovascular	1.088 (1.078,1.098)	1.016 (1.005,1.028)
IHD	1.126 (1.112,1.140)	1.027 (1.011,1.043)
CBV	1.126 (1.103,1.150)	1.057 (1.029,1.085)
CHF	0.986 (0.953,1.021)	0.970 (0.930,1.012)
All Respiratory	1.056 (1.038,1.074)	0.989 (0.968,1.010)
COPD	1.023 (0.999,1.047)	0.998 (0.969,1.027)
Pneumonia	1.078 (1.044,1.114)	0.943 (0.905,0.982)
All Cancer	1.025 (1.012,1.038)	1.016 (1.001,1.031)
Lung cancer	0.995 (0.972,1.018)	0.986 (0.958,1.014)

[1]Wang et al. (2020)

Bayesian Kernal Machine Regression (BKMR)
(Bobb et al., 2018)

$$Y_i = h(z_{i1}, \dots, z_{iM}) + x_i'\beta + \epsilon_i$$

can be non-linear, non-additive

where Y_i is response for individual i ($i = 1, \dots, n$), z_{im} the m^{th} components, h the unknown exposure-response function to be estimated, β the effect of covariates

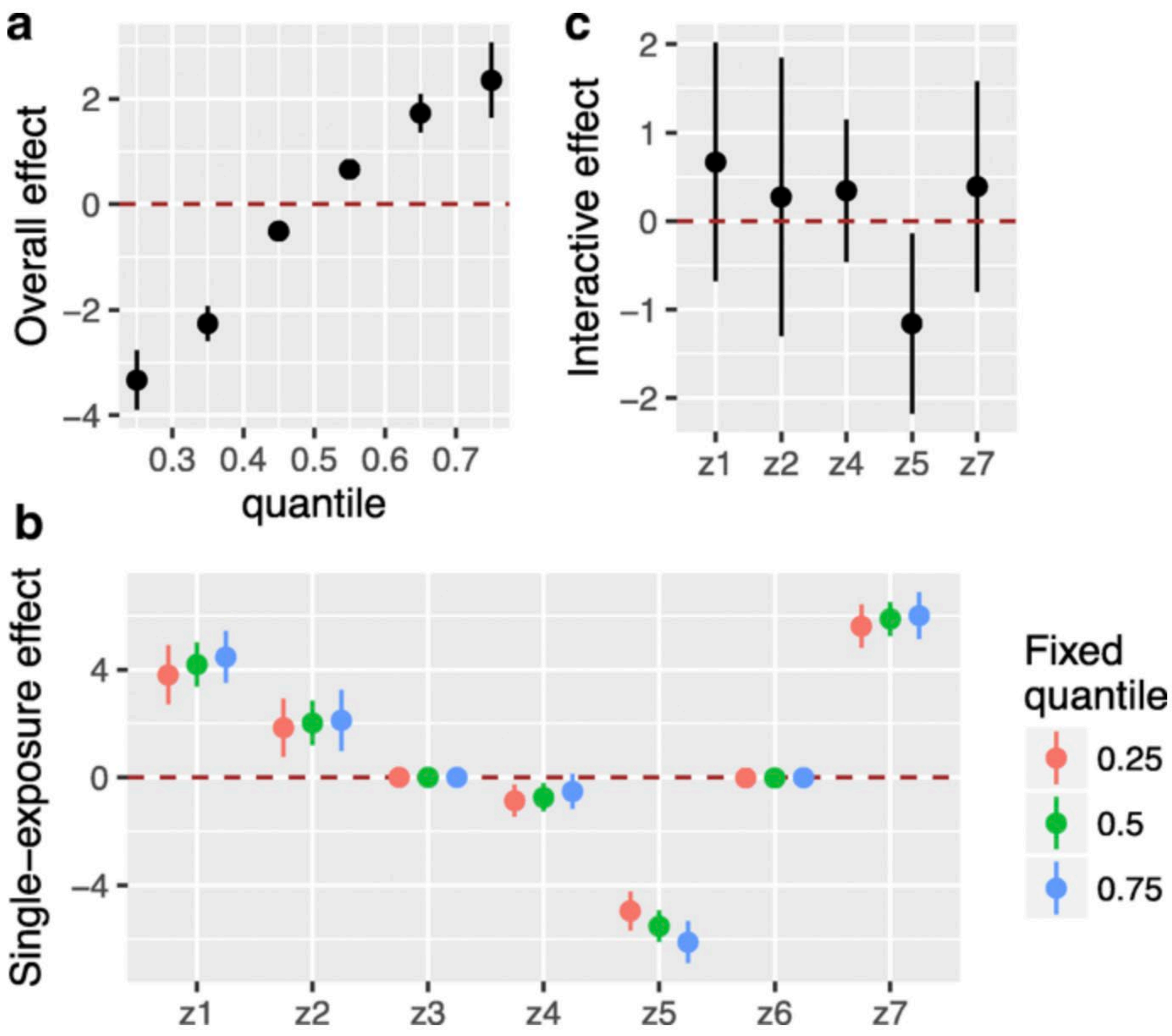
- Assesses mixture impacts **flexibly**, accounting for varying correlations amongst components
- **Validated** using a simulated dataset
- Can be **implemented** using open-source R software package (*bkmr*), allowing
  - Visualization of the exposure-response relationship
  - Summaries of overall, single-component, interactive health effects.
  - Time efficient analysis

BKMR

Bobb et al. (2018)

Simulated data set, continuous outcome

Fig. 3



Bobb et al. (2018), *Biostatistics*

Summary

- Evidence of differential toxicity by PM component → need more studies (esp. long-term) to assess or establish consistency
- Growing number of new resources and methods for these studies
  - For single component, source models, advances will result from availability of more data for more PM components, with finer spatial, temporal resolution
    - Issues of exposure error, correlations will still be of concern
    - Useful for regulation, hypothesis generation, interpretation of results from mixture methods
  - To assess mixture effects, clearly need approaches that simultaneously and flexibly consider multiple components
    - Residual-based methods, but interpretation of coefficients tricky
    - Need more studies using BKMR, other methods, including those applied to same data sets