THE UNIVERSITY of TENNESSEE KNOXVILLE

19th ANNUAL CMAS Conference
Oct. 26-30 | Virtual

Assessing Heterogeneity of the Burden of Disease of $PM_{2.5}$ Exposure at Diverse Urbanization Levels with CMAQ-fused Data

Cheng-Pin Kuo [a], Joshua S. Fu [a]

[a] Department of Civil and Environmental Engineering, University of Tennessee Knoxville, Knoxville, TN, USA

October 29, 2020

Burden of Disease

“ *Burden of disease is concept that describes death and loss of health due to diseases, injuries and risk factors.* ”

World Health Organization (WHO)

“ *Global Burden of Disease (GBD) estimated that particulate matter pollution caused 2.94 million deaths during 2017.* ”

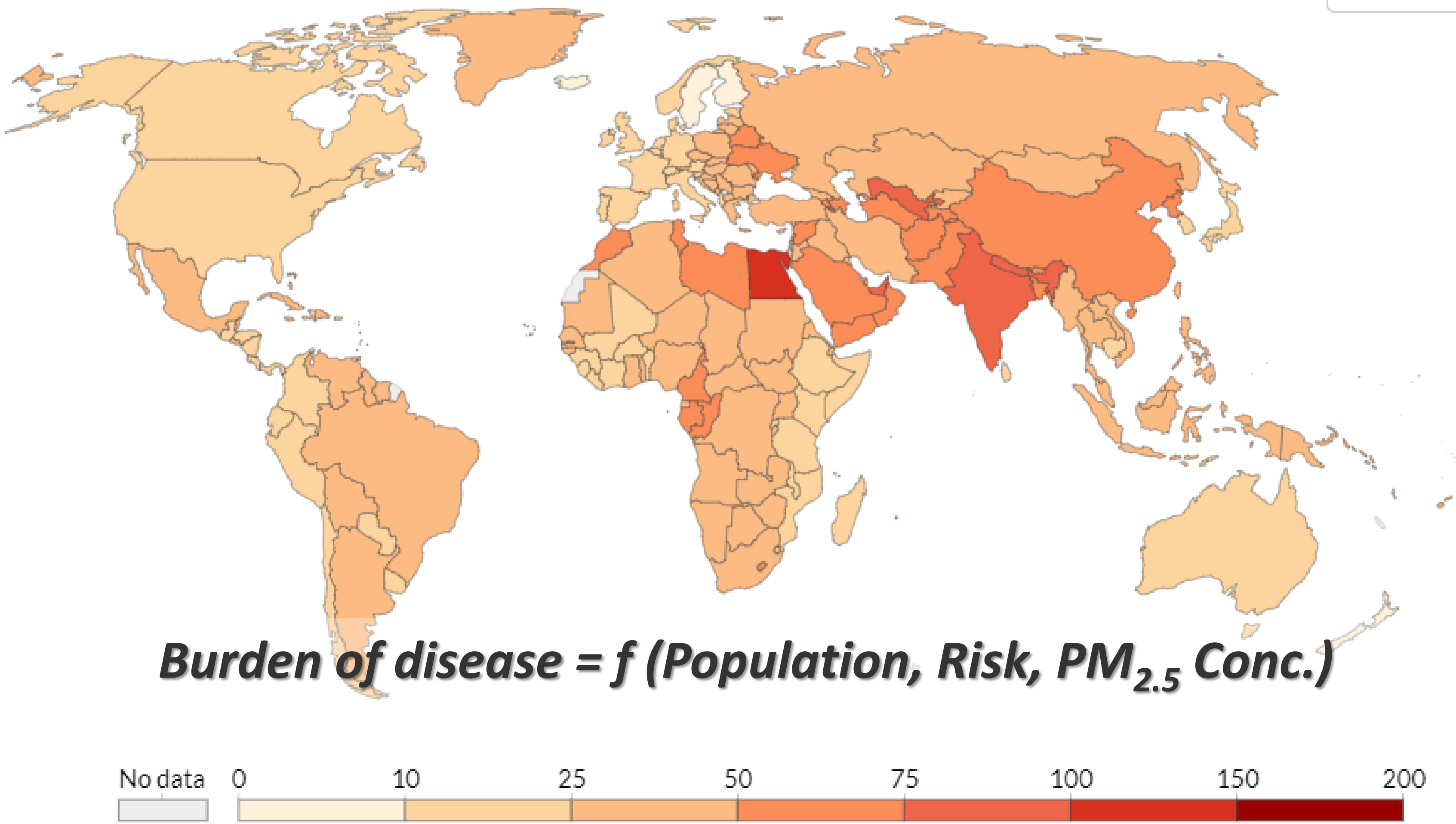
(Stanaway et al., 2018)

Outdoor air pollution death rate, 2017

The number of deaths attributed to outdoor ozone and particulate matter pollution per 100,000.

Our World in Data

World

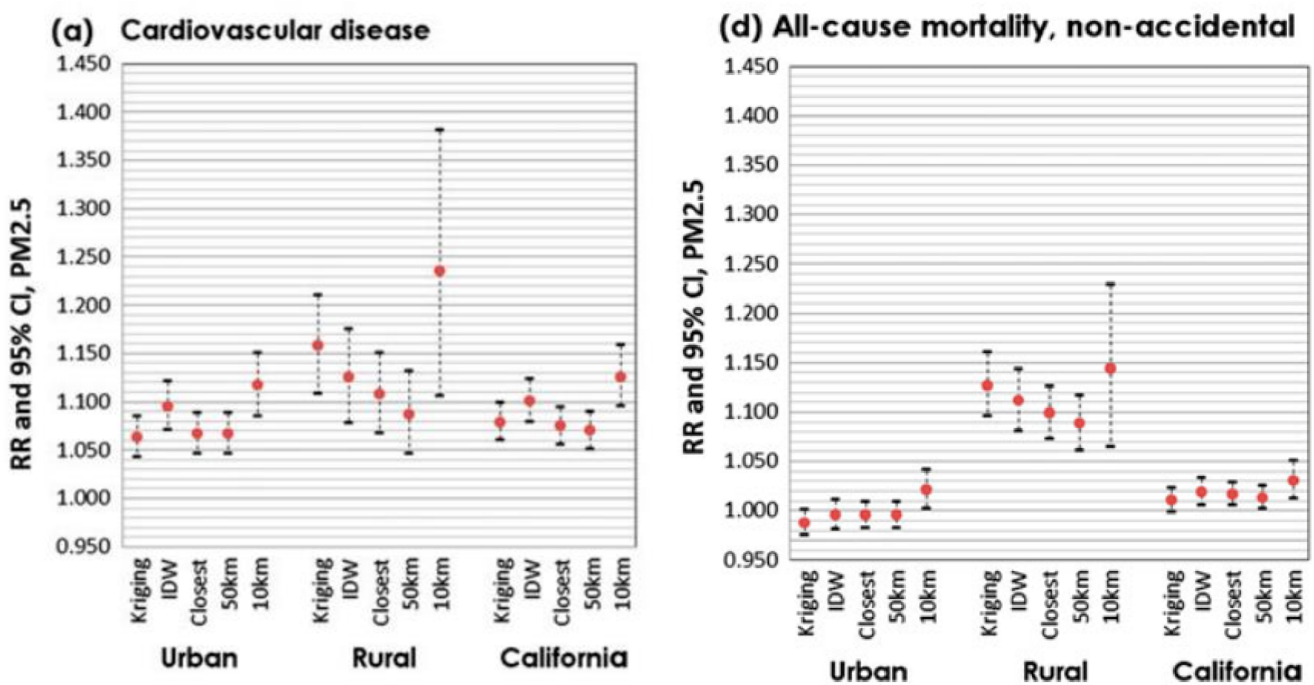


Burden of disease = f (Population, Risk, $PM_{2.5}$ Conc.)

IHME, Global Burden of Disease, https://ourworldindata.org/outdoor-air-pollution

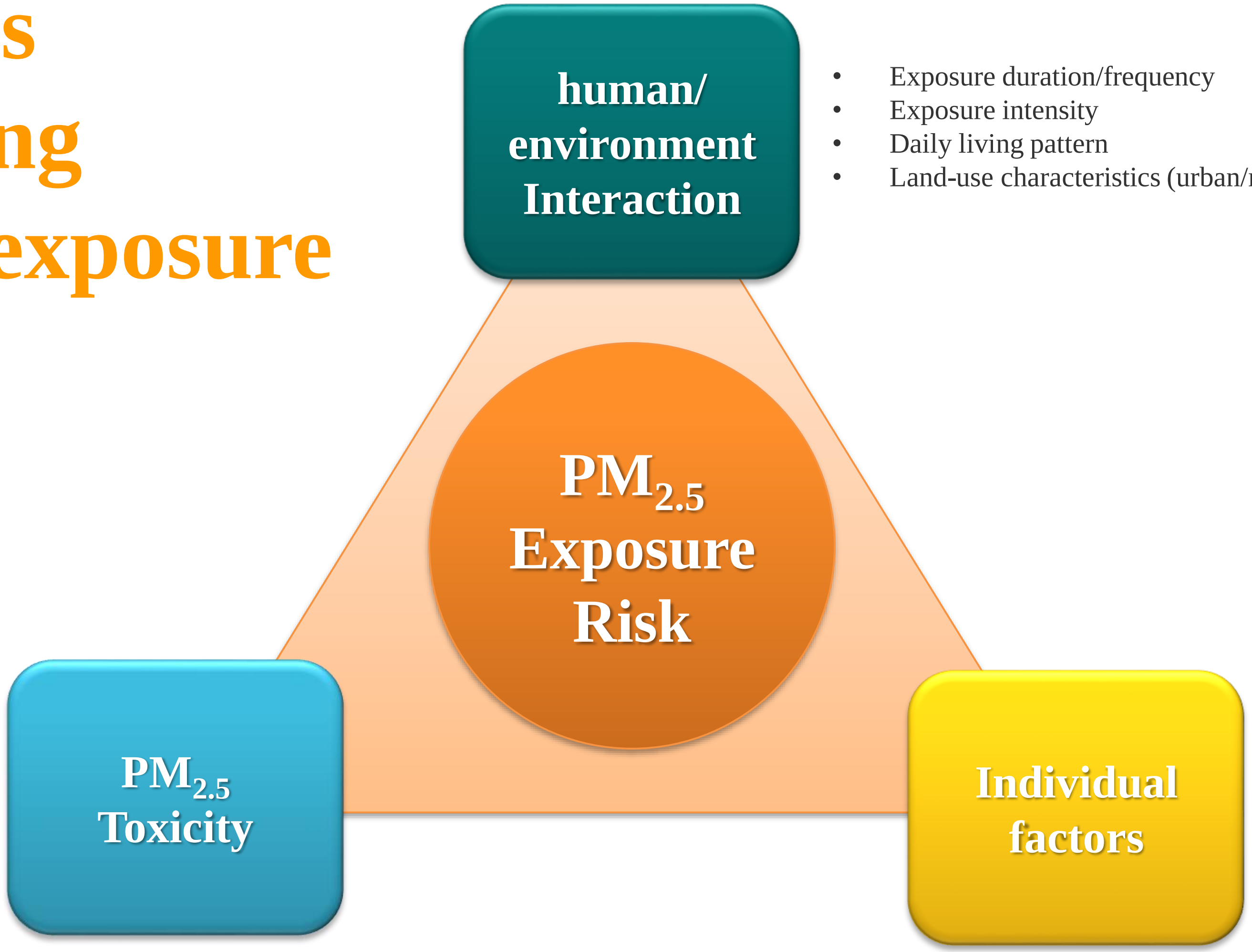
Motivations

- Although GBD considered risk of $PM_{2.5}$ exposure at national-wide level, different $PM_{2.5}$ exposure risks at finer geological scales such as urban and rural areas had been identified in previous studies. *(Garcia et al., 2016; McGuinn et al., 2017)*



Source: Garcia et al., 2016

Factors affecting $PM_{2.5}$ exposure risk



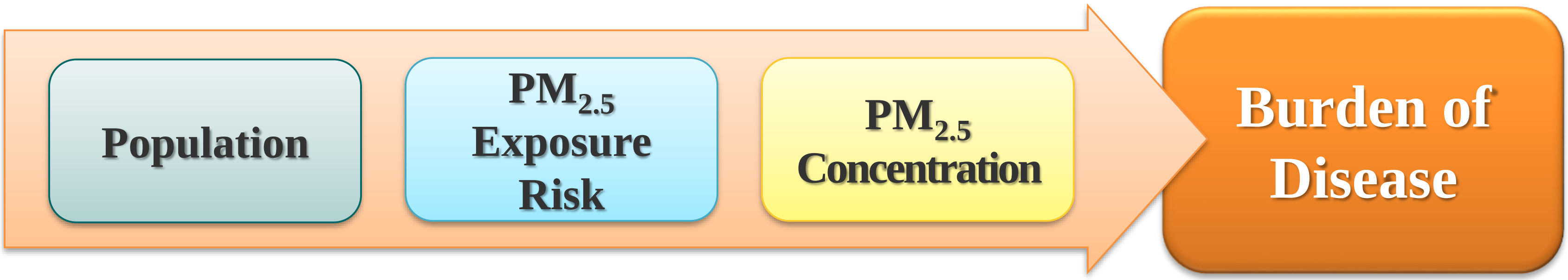
- Exposure duration/frequency
- Exposure intensity
- Daily living pattern
- Land-use characteristics (urban/rural)

- $PM_{2.5}$ compositions
- nitrate, sulfate, ammonium, OC, EC
- diesel particles, heavy metals, dioxin, other HAPs

- Vulnerability (gender, age, gene)
- Knowledge
- Social-economic status
- Medical accessibility

Uncertainty of BD estimation

Burden of disease = f (Population, Risk, $PM_{2.5}$ Conc.)



- Different sources of risk values and ambient $PM_{2.5}$ concentration would pose uncertainty for burden of disease estimation
  - Risk values: USEPA recommended values, nation-specific values, finer-scale values
  - Ambient concentration data: monitoring data, modeled data
- In this study, the uncertainty between different methods to estimate BD should be quantified as well.

Objectives

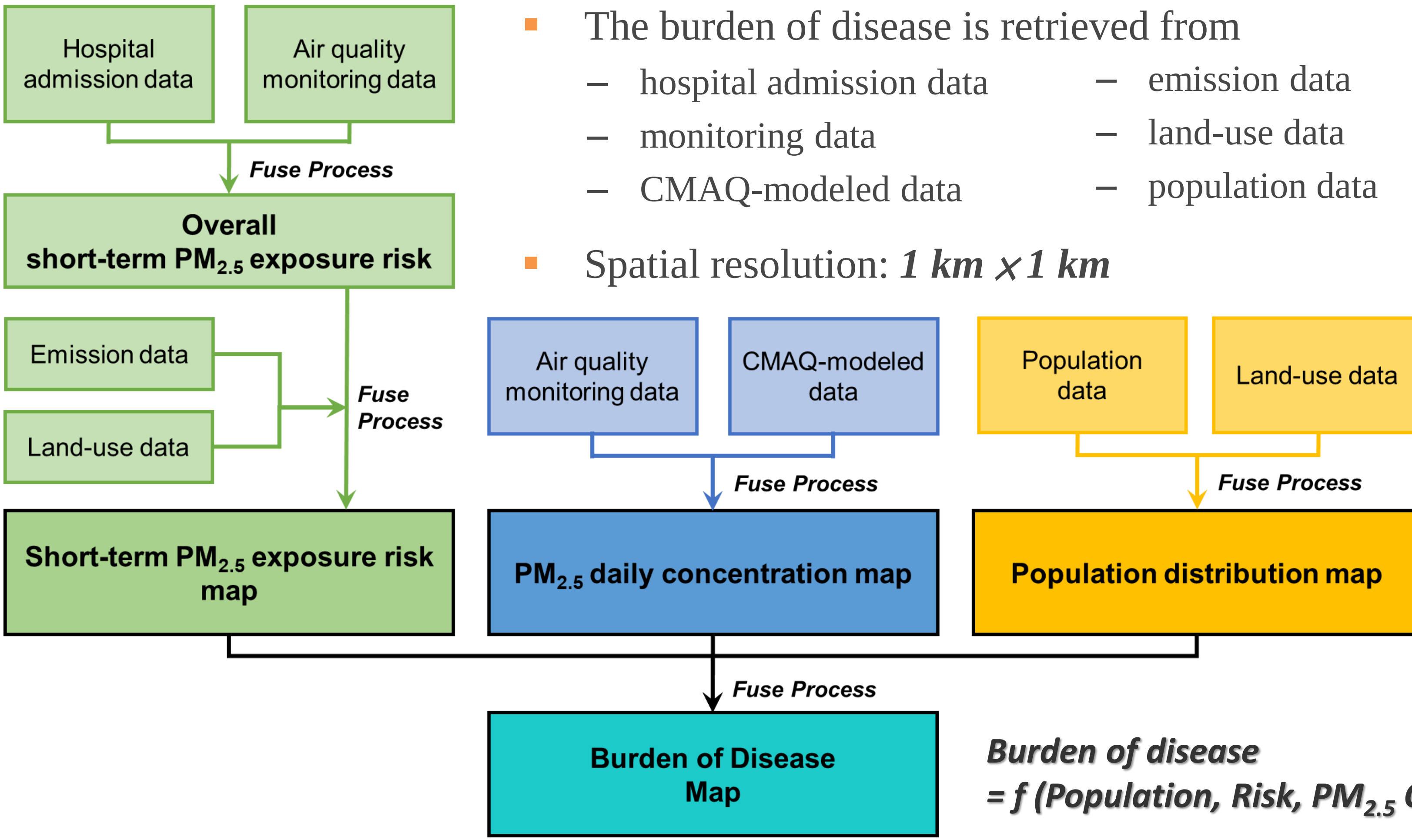
- Develop a feasible and generalizable burden of disease estimation approach that can consider $PM_{2.5}$ exposure risks in diverse urbanization levels.
- Quantify the uncertainty of present burden of disease estimation methodology with different risk values and $PM_{2.5}$ exposure data

Study region

- Population:
  - about 1.9 millions
- Area
  - about 2,200 km^2
- Population density
  - 856 people/km^2

Methods

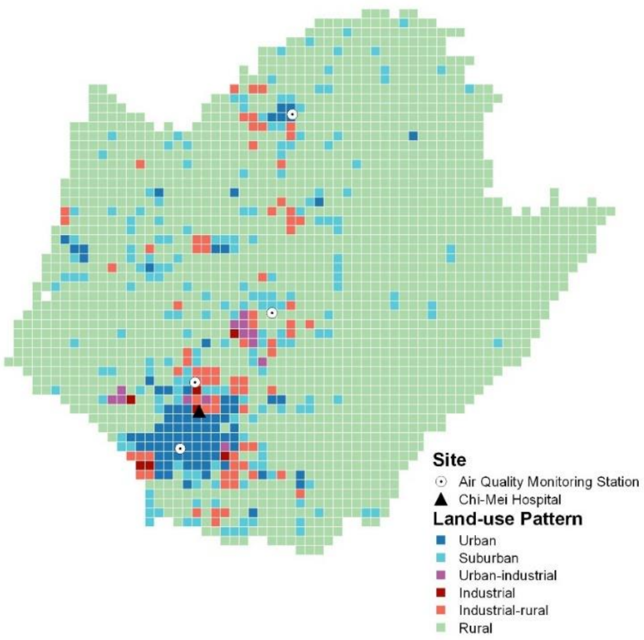
- The burden of disease is retrieved from
  - hospital admission data
  - monitoring data
  - CMAQ-modeled data
  - emission data
  - land-use data
  - population data
- Spatial resolution: ***1 km × 1 km***



Burden of disease
= f (Population, Risk, $PM_{2.5}$ Conc.)

Methods

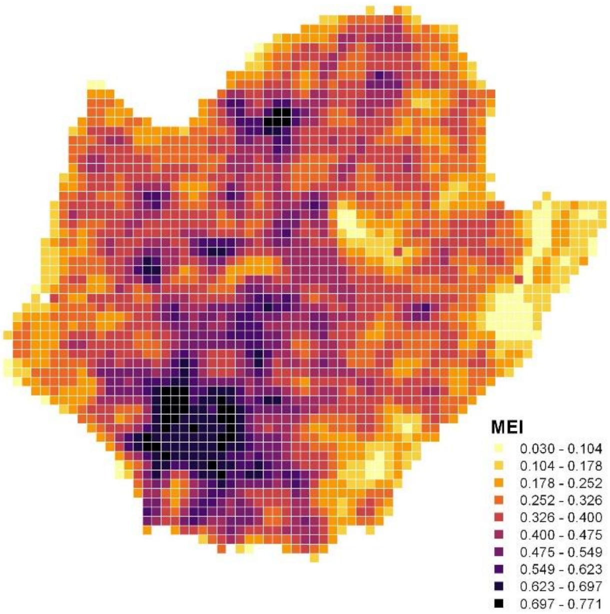
- Hospital admission data (2006-2016)
  - 12,524 subjects with cardiovascular diseases (CVDs)
  - 37,846 non-accidental deaths
- Air quality monitoring data (4 sites, 2006-2016)
  - averages of 24 hr before admission were calculated
  - $PM_{2.5}$, PM_{10}, NO, SO_2, CO, O_3
  - ambient temperature, relative humidity, wind speed
- CMAQ-modeled data in 2013
  - CMAQ v5.2 / WRF
  - Taiwan Emission Data System (TEDS) version 9.0
  - CMAQ-modeled data was further fused with 2013 monitoring data



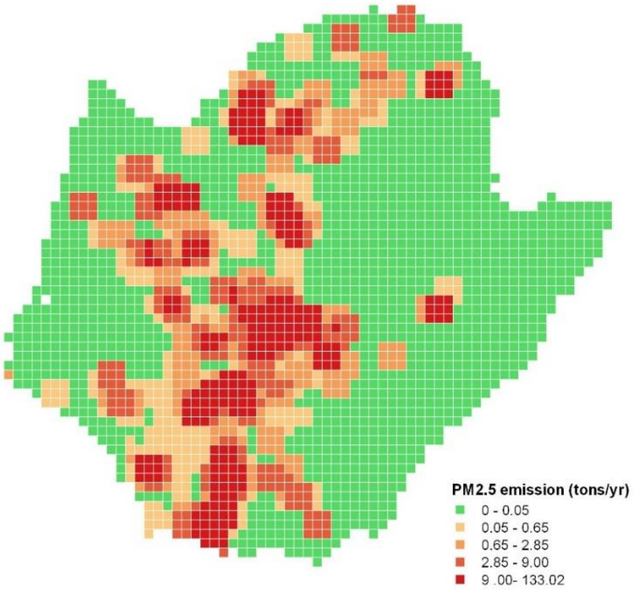
Methods

- Urbanization level was defined by **Mean Entropy Index (MEI)** and **$PM_{2.5}$ emission density**
  - **MEI** represents the diversity and heterogeneity of land-use patterns in each grid
  - Equally divided subjects to 3 groups and 3*3=9 categories could be obtained

(a) Mean Entropy Index (MEI)

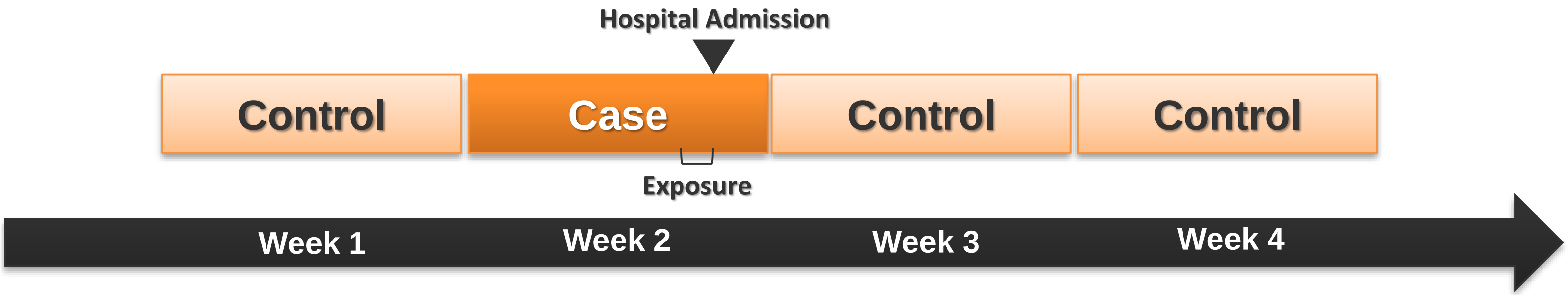


(b) $PM_{2.5}$ emission density



Case-crossover study design and Stratified analysis

- Case-crossover study design
  - to investigate the relationship between $PM_{2.5}$ exposure and health outcomes
  - used subjects themselves as controls (**self-matching**) but at different time period
  - Conditional logistic regression



- Stratified analysis
  - The same regression model was applied to 9 categorized groups
  - Different risks (odds ratios, ORs) for each groups could be retrieved

Uncertainty analysis

- Burden of disease estimation: **concentration-response functions (CRFs)**
  - quantify the increased hospital admissions due to short-term $PM_{2.5}$ exposure

$$\mathbf{Y} = \mathbf{E_0} \times \mathbf{Pop} \times \left(1 - e^{-\beta \times (\mathbf{C} - \mathbf{C_0})}\right) \times \mathbf{A}$$

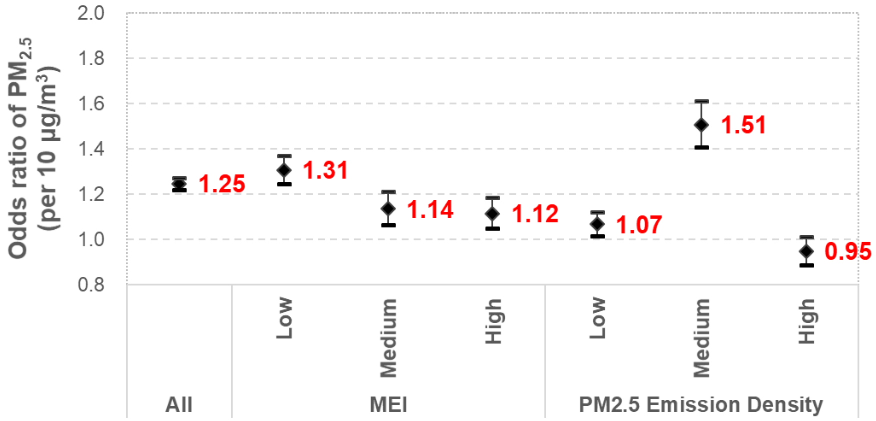
- Uncertainty under different scenarios
  - Risk parameters (β):
    - USEPA recommended risk
    - Average risk
    - Heterogeneously distributed risks
  - Daily exposure data (**C**):
    - Monitoring data
    - CMAQ-fused data

$$\textbf{Fused PM}_{\mathbf{2.5}\,d,g} = \frac{\textbf{Obseved PM}_{\mathbf{2.5}\,d,s}}{\textbf{Modeled PM}_{\mathbf{2.5}\,d,s}} \times \textbf{Modeled PM}_{\mathbf{2.5}\,d,g}$$

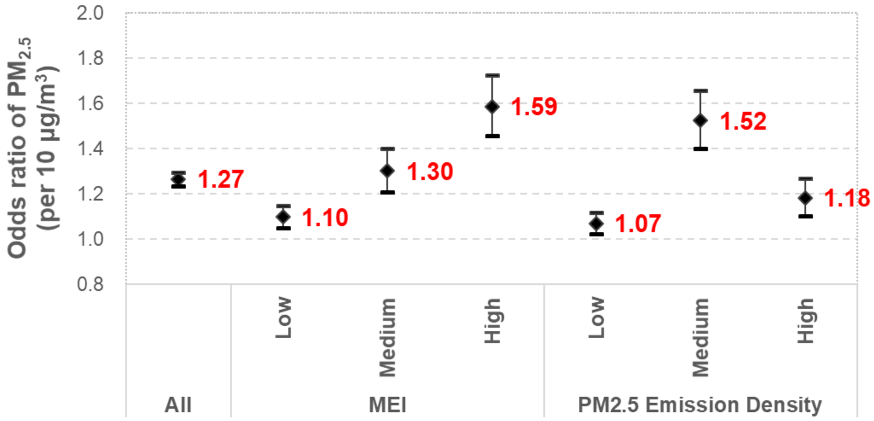
Results

- Overall risks (ORs) was
  - 1.25 (95% CI: 1.22-1.27) for death
  - 1.27 (95% CI: 1.24-1.30) for CVDs
- The risks varied with MEI levels and neighboring $PM_{2.5}$ emission densities. With the same $PM_{2.5}$ exposure level
  - living in low MEI areas would have high risk of death
  - living in high MEI areas would have high risk of CVDs
  - neighboring to medium-level $PM_{2.5}$ emission density would have higher risk.

(a) Death



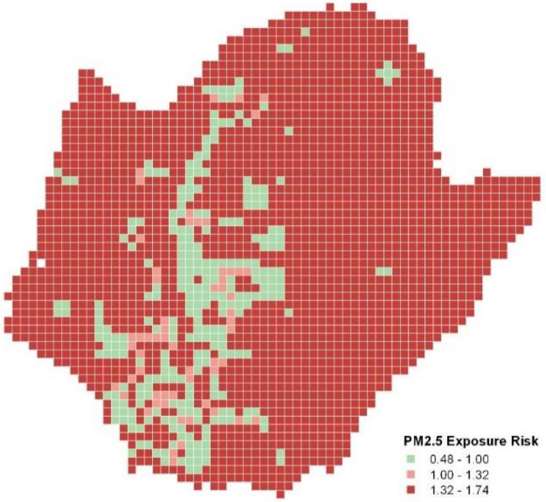
(b) Cardiovascular Disease



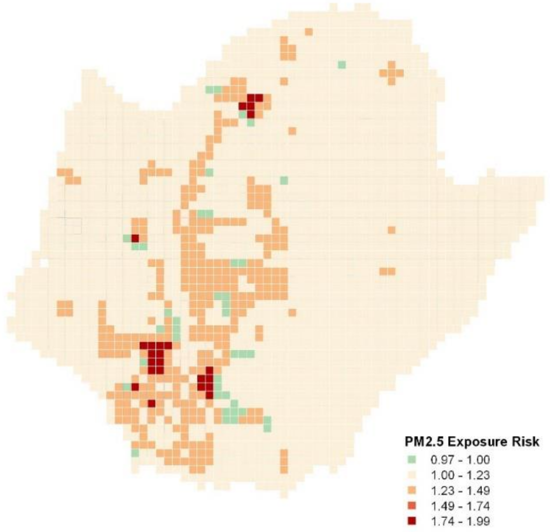
Results – Risk Map

- When applying categorized risks to the map,
  - Death: higher risk in urban and rural, lower in suburban
  - CVDs: higher in urban and suburban, lower in rural
- Different patterns for death and CVDs implied the importance of medical accessibility
  - CVD patients in urban could seek medical treatment more easily to avoid death

(a) Death



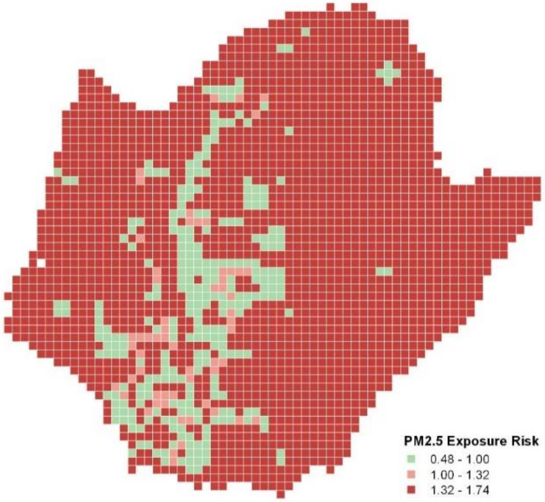
(b) Cardiovascular Disease



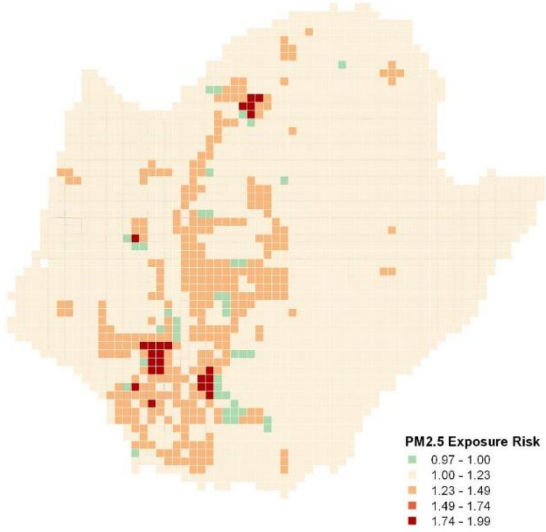
Results – Risk Map

- Rural: “**non-metropolitan penalty**”
  - lower accessibility of medical resources
  - financial consideration
  - individual knowledge and awareness
  - frequent outdoor activities
- Urban
  - high accessibility of medical resources
  - higher awareness
  - higher social-economic status
  - higher $PM_{2.5}$ at breathing level

(a) Death



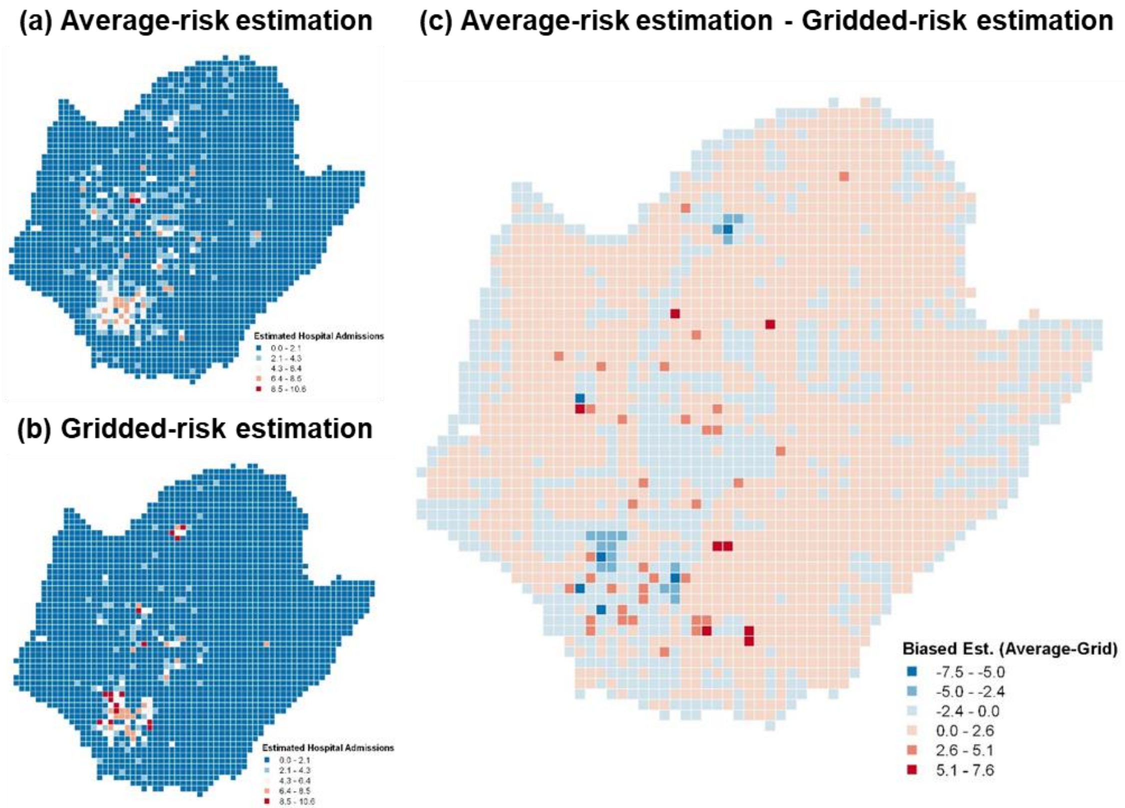
(b) Cardiovascular Disease



Uncertainty Analysis

- To illustrate the spatial difference, use CVD admissions map as an example

Burden of disease = f (Population, Risk, $PM_{2.5}$ Conc.)



Uncertainty Analysis

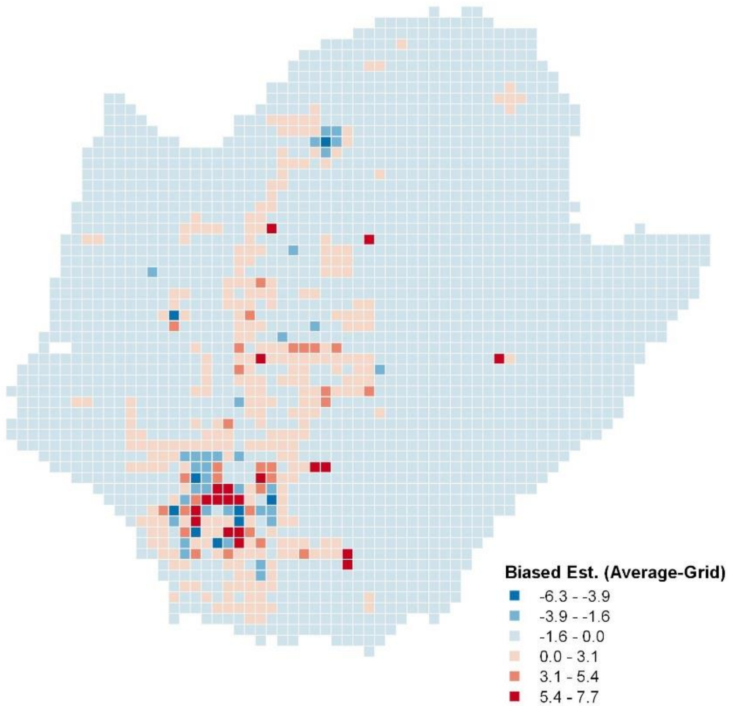
Scenario	Applied Risk	N (%)		
		Monitoring data	CMAQ-Fused data	Difference
Death				
Gridded Local Risk	1.000 - 1.742	1,699 (100%)	2,137 (100%)	-437 (-20%)
Averaged Local Risk	1.245	1,698 (100%)	2,266 (106%)	-568 (-25%)
USEPA Recommended Risk	1.170	1,309 (77%)	1,790 (84%)	-481 (-27%)
Cardiovascular Disease				
Gridded Local Risk	1.000 - 1.992	1,297 (100%)	1,750 (100%)	-453 (-26%)
Averaged Local Risk	1.266	1,790 (138%)	2,375 (136%)	-585 (-25%)
USEPA Recommended Risk	1.002	178 (14%)	262 (15%)	-85 (-32%)

- Using monitoring data would underestimate 2013 hospital admissions by 20%-32% compared with CMAQ-fused data.
- Compared with using gridded risk
  - Using USEPA recommended risk would largely underestimate by 23%-86%
  - Using averaged local risk would overestimate by 0%-38%

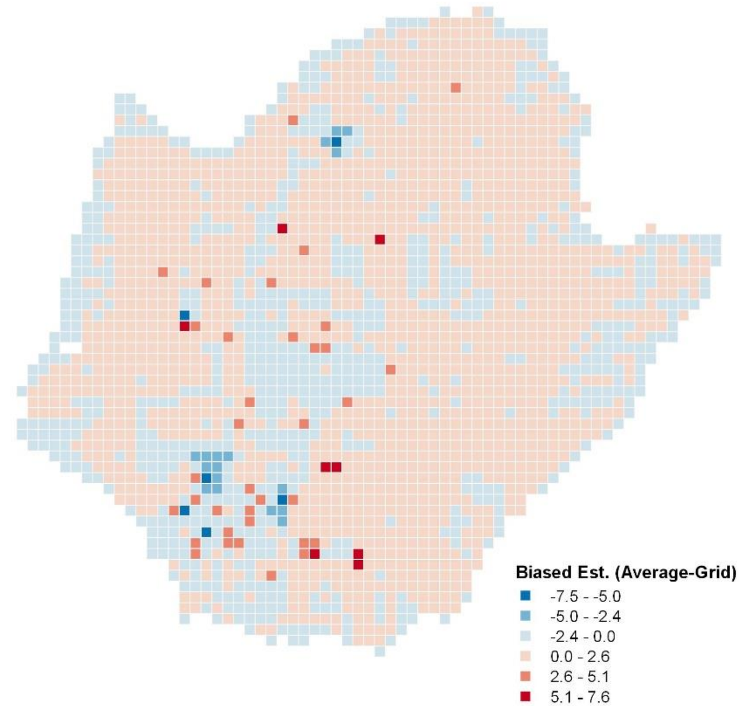
Uncertainty Analysis

- Compared with using gridded-risks, using averaged risks would
  - Death: underestimate in rural and overestimate in suburban and urban
  - CVD: underestimate in urban and suburban and overestimate in rural

(a) Death



(b) Cardiovascular Disease



Conclusions

- The heterogeneity of land-use patterns (MEI) and neighboring emission would affect the vulnerability of $PM_{2.5}$ short-term exposure, and the spatial distribution depend on health outcomes
  - The higher death risk occurs in rural but the higher CVD risk occurs in urban
- The burden of disease estimation would be largely biased by risk parameters and $PM_{2.5}$ exposure data.
  - Using monitoring data and ignoring the spatial heterogeneity of risks could have a considerable bias (-32% ~ 85%)
- Our evaluation methodology could be generalized to any city or country. Cities or countries without appropriate risk estimates could utilize our approach to develop their own burden of disease estimation.

Thanks for your attention!

Q&A

cmasconference2020@unc.edu