Examine the Association between Temperature Variation and Cardiovascular Disease Hospitalization

A Case Study in Shanxi, China

Miao Cai

Department of Epidemiology & Biostatistics,
Saint Louis University

2019-05-01

Introduction

Theoretical background

Global climate change

- Frequent heat shocks and cold spells[1,2]
- 12.6 M deaths are attributable to climate change[3]

CVD

- Cardiovascular disease (CVD) is the leading cause of death worldwide and in China
- 15.2 M killed by CVD, which accoutns for 26.7% of all deaths[4]

Research Question

Is **temperature variation** associated with **CVD hospitalizations**?

- biological association
- research gap: no epidemiological evidence

Data source

Shanxi Province, China

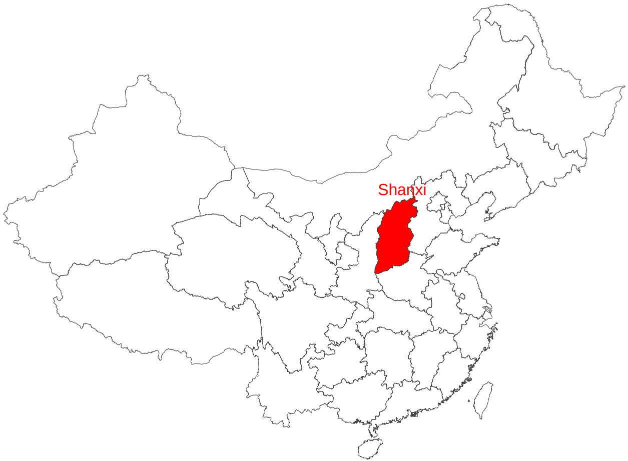


Figure 1: The location of Shanxi Province in China

Shanxi inpatient database

CVD inpatients

- 175 hospitals
- 95% of administrative districts (ADs)
- $\sim$ 1.2 million CVD patients
- 2015 - 2017

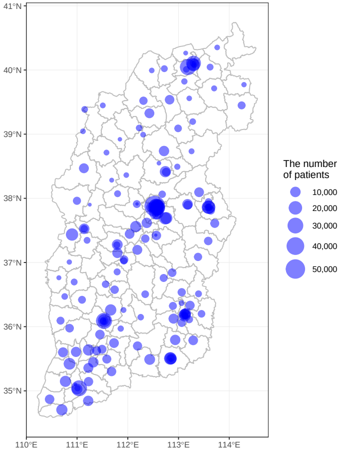


Figure 2: Geographic distribution of 175 hospitals in Shanxi, China

The number of CVD patients by month

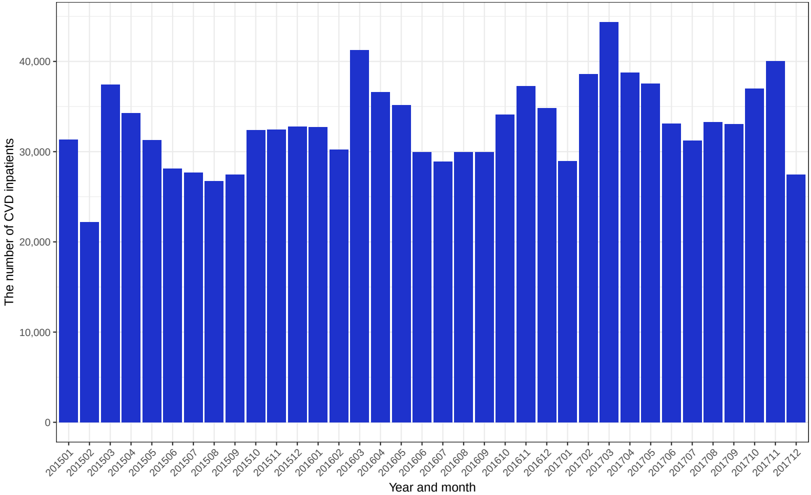
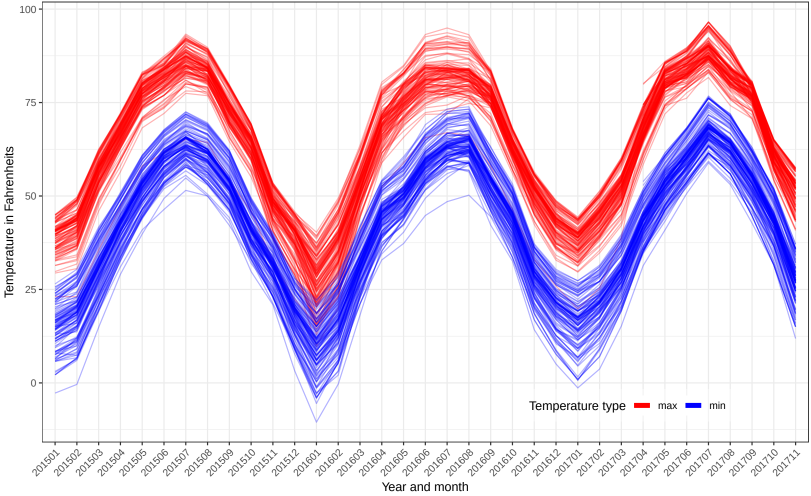


Figure 3: The number of CVD patients by year and month

Temperature variation definition

- Vertical variation: $V_v = T_{max} - T_{min}$
- Horizontal variation: $V_h = \sum_{i=\mathrm{d}-7}^{\mathrm{d}} \left| T_i - T_{i-1} \right|$



Temperature variation trend

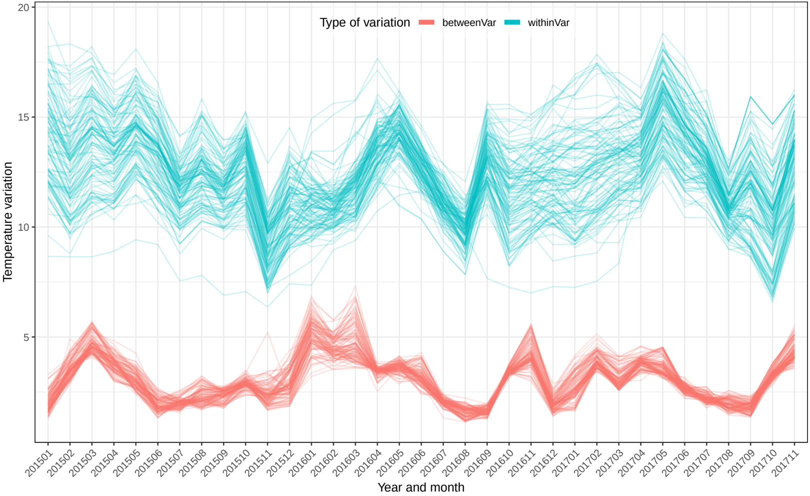


Figure 5: Temperature variation trend in 118 ADs in Shanxi, 2015 - 2017

Shanxi Statistical Yearbook

County/city level

- 118 ADs

Variables available

- Population
- Gender
- Rural
- GDP
- Others

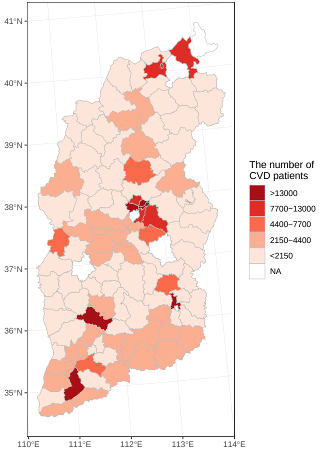


Figure 6: Choropleth map patient distribution in each county and city in Shanxi, China

Final sample

- 109 ADs
- 12 months in 3 years
- $109 * 12 * 3 = 3,924$ rows

Data cleaning, visualization, statistical models, and reporting are conducted in `R 3.6.0`.

Statistical models

Model selection

The outcome is the count of CVD patients:

- Must be non-negative,
- integers,
- highly righ-skewed.

OLS assumpts are **not met**!

Poisson regression is a natural model for count data. We consider:

1. non-spatial Poisson regression
2. random effects Poisson regression
3. spatial Poisson regressions

Non-spatial Poisson model

$$Y_i \sim \text{Poisson}(\text{Pop}_i * \lambda_i)$$
$$\log(\lambda_i) = \beta_0 + \sum_k \beta \mathbf{X}$$

- Y_i: The number of CVD inpatients
- Pop_i: The total population as an offset
- λ_i: The rate parameter of Poisson distribution, both the mean and vaiance
- Predictor variables: temperature variation (horizontal and vertical), average temperature, log of GDP per capita, the percent of female, and the percent of rural population

Non-spatial Hierarchical Poisson model

Assuming that each ADs have its own characteristics, and observations are **conditionally independent**.

$$
\begin{aligned}
Y_{ij} &\sim \text{Poisson}(\text{Pop}_{ij} * \lambda_{ij}) \\
\log(\lambda_{ij}) &= \beta_{0j} + \beta_{1j} HV_{ij} + \beta_{2j} VV_{ij} + \sum_{k-2} \beta \mathbf{X} + u_{0j} \\
\beta_{0j} &\sim N(\mu_0, \sigma_0^2) \\
\beta_{01} &\sim N(\mu_1, \sigma_1^2) \\
\beta_{02} &\sim N(\mu_2, \sigma_2^2)
\end{aligned}
$$

- Random intercepts: β_{0j} for each AD j
- Random slopes: β_{1j} and β_{2j} for each AD j

Spatial Poisson models

Spatial lag Poisson model

- An autoregressive parameter ρ
- A weighting matrix $\mathbf{W}$

$$Y_i \sim \text{Poisson}(\text{Pop}_i * \lambda_i)$$
$$\log(\lambda_i) = \beta_0 + \sum_k \beta \mathbf{X} + \rho \mathbf{W} y$$

Geographically weighted Poisson model

Still at the initial proof of concept stage. It can be conducted in the `R` package `spgwr` by Roger Bivand (2017).

Results

Moran's I - CVD patients

```r
coords = hloca %>%
  select(longitude, latitude) %>%
  as.matrix()
nb = dnearneigh(coords,d1=0, d2 = 70, longlat = TRUE)
moran.test(hloca$N, nb2listw(nb, style="W"))
```

```
##
##  Moran I test under randomisation
##
## data:  hloca$N
## weights: nb2listw(nb, style = "W")
##
## Moran I statistic standard deviate = 2.4915, p-value = 0.00636
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic       Expectation          Variance
##       0.075114045      -0.006896552       0.001083440
```

non-spatial Poisson models

Table 1: Parameter estimates of non-spatial Poisson models

	Dependent variable:	
	Poisson	Mixed-effects Poisson
	(1)	(2)
Between variance	0.072*** (0.001)	0.065*** (0.006)
Within variance	−0.0003 (0.001)	0.015*** (0.004)
Mean temperature	−0.001*** (0.0001)	−0.002*** (0.0001)
Log GDP	0.237*** (0.002)	0.425*** (0.016)
Rural	−1.022*** (0.007)	−1.037*** (0.141)
Female	0.129*** (0.001)	0.011*** (0.004)
Constant	−13.347*** (0.065)	−8.114*** (0.224)
Log Likelihood	-141,512.000	-29,727.990
Akaike Inf. Crit.	283,038.000	59,481.970
Note:		*p<0.1; **p<0.05; ***p<0.01

Spatial Poisson models

To be added

- Spatial lag Poisson model
- Geographically weighted Poisson model

Discussion

Limitations

- Assumptions: the patients did not seek care in other ADs
- Omitted variable bias: air pollution, older population
- Different weights for different severity of CVD hospitalization?

Future directions

- Add time series analysis in several major cities?
- Different effects in cold and hot days?
- Hierarchical spatial Poisson regression?

References

1 Ryti NR, Guo Y, Jaakkola JJ. Global association of cold spells and adverse health effects: A systematic review and meta-analysis. *Environmental health perspectives* 2015; **124**: 12–22.

2 Huber D, Gulledge J. The global links between extreme weather and climate change. *Extreme Weather Events* 2017;: 17.

3 Watts N, Adger WN, Ayeb-Karlsson S *et al.* The lancet countdown: Tracking progress on health and climate change. *The Lancet* 2017; **389**: 1151–64.

4 The World Health Organization. The top 10 causes of death. 2018.