# The Effect of Hospital Networks on Clostridium difficile Infection Rates

Jordan Baker, Serenity Budd, Carissa Comnick Advisor: Dr. Daniel Sewell

#### What is C. diff?

- Clostridium difficile, bacteria
- Symptoms
  - Colitis
  - Diarrhea
  - $\circ$  Nausea
  - Fever
  - Dehydration (can lead to kidney failure)
  - Death
- How it spreads
  - $\circ$  ~ Spores passed in feces when infected don't wash hands; spread to surfaces
  - $\circ$   $\quad$  Touching contaminated surfaces and then touching mouth, nose, or eyes

#### What is C. diff?

- Risk factors
  - $\circ \quad \text{Advanced age} \quad$
  - $\circ \quad \ \ {\rm Prior \ antimic robial \ use}$
  - $\circ$  Severity of illness
  - $\circ \quad \text{Length of stay} \\$
  - $\circ$  Being in a hospital with more patients
  - Being in a network of hospitals with more patients
- How it is treated
  - Antibiotics
  - Surgery
  - $\circ$  Fecal microbiota transplant
  - Probiotics

#### Research Question

How much does the network of patient sharing among hospitals contribute to C. diff rates?



#### Hospital Network Data

- Hospital data C. diff infections
  - $\circ$  410 California hospitals identified by AHAID
  - Monthly counts (2008-2011) for number of primary and secondary C. diff cases
  - $\circ \quad \ \ {\rm Average\ monthly\ length\ of\ stay}$
  - $\circ$  Monthly percent of patients over 65
  - $\circ$  Other demographic variables
- Hospital data locations
  - $\circ$  Data from hospitals nationwide
  - $\circ \quad {\rm AHAID \ and \ latitude/longitude \ of \ each \ hospital}$
- Transfer list
  - $\circ$  Transfers from one hospital to another
  - Identified by Department of State Hospital ID

## Data Cleaning

- Removed missing data and hospitals not in both lists
- Fixed repeated months due to long term patients
- Fixed repeated hospital locations
- Converted the transfer list to an adjacency matrix
  - Did not use patients readmitted to the same hospital
- Final data:
  - 388 hospitals
  - 306,139 transfers
  - $\circ$  ~ CDI mean monthly ranged from 0 to 72 cases
  - $\circ$  Mean monthly length of stay ranged from 1.3 to 76.5 days
  - $\circ$   $\,$  Mean monthly % over 65 ranged from 0 to 93 percent  $\,$
  - $\circ$   $\,$  Mean monthly admissions ranged from 3 to 4619 days  $\,$





Model - Linear Network Autocorrelation Model  $y_i = X'_i \underline{\beta} + \rho \sum_{j \neq i} A_{ij} y_j + \epsilon_i, \ \epsilon_i \sim N(0, \sigma^2)$  $y_i$  is the estimated average monthly number of C. diff cases for hospital i

 $X_i$  is vector of total admissions, % over 65, length of stay for hospital i

 $\underline{\beta}$  is a vector of covariate regression coefficients

 $\boldsymbol{\rho}$  is the network effect coefficient

 $\sum A_{ij} y_j$  is every patient transferred from each j to i, multiplied each by the amount of C. diff at hospital j $\epsilon_i$  is a value of the random error Model - Linear Network Autocorrelation Model  $\underline{y} = X\underline{\beta} + \rho A\underline{y} + \underline{\epsilon}, \ \underline{\epsilon} \sim N(\underline{0}, \sigma^2)$ 

Multivariate normal distribution:

$$\underline{y} \sim N((I - \rho A)^{-1} X \underline{\beta}, \sigma^2 (I - \rho A)^{-1} (I - \rho A')^{-1})$$

#### Model - Add in Spatial Component

• This is the statistical model that we used

$$\underline{y} = X\underline{\beta} + \rho A\underline{y} + \underline{\epsilon}, \ \underline{\epsilon} \sim N(\underline{0}, \sigma^2 \Sigma_{\phi})$$
$$[\Sigma_{\phi}]_{ij} = e^{-\phi d_{ij}}$$
$$\underline{y} \sim N((I - \rho A)^{-1} X\underline{\beta}, \sigma^2 (I - \rho A)^{-1} \Sigma_{\phi} (I - \rho A')^{-1})$$

- $d_{ii}$  distance between hospitals using longitude & latitude
- $\phi$  spatial parameter

#### Results - 95% CI for Model Coefficients

	Estimate	Lower Bound	Upper Bound
Intercept	-4.35	-5.39	-3.31
Mean Length of Stay	0.03	-0.02	0.08
Mean % Over 65	9.14	7.0	11.28
Mean Admissions	0.009	0.008	0.01
ρ	0.00007	0.00004	.0001
$\sigma^2$	184.9	137.0	244.7
$\log(\text{Spatial } \phi)$	5.0	-24.81	34.82

#### **Results - Counterfactual Research Question**

How much would CDI rates change if there were no network effect?

$$\frac{E(\underline{y'}\underline{1}|\rho=0) - E(\underline{y'}\underline{1}|\rho\neq 0)}{E(\underline{y'}\underline{1}|\rho\neq 0)} = \frac{(X\underline{\beta})'\underline{1} - ((I-\rho A)^{-1}X\underline{\beta})'\underline{1}}{((I-\rho A)^{-1}X\underline{\beta})'\underline{1}}$$

$$= -.089 \rightarrow -9\%$$

 $\sim$  3,000 fewer cases per year

 $\sim$  \$16.5 million less spent on C. diff patient healthcare costs per year

#### Future Research

- Leverage network topology to inform intervention and surveillance strategies
- Network effect on other healthcare-associated infections
- Treat the data as a time series instead of aggregating over several years



#### References

- Centers for Disease Control and Prevention. (2015). *Medical illustration of Clostridium difficile* [graphic]. Retrieved from https://www.cdc.gov/hai/organisms/cdiff/cdiff-patient.html
- Doreian, P., Teuter, K., & Wang, C. (1984). Network autocorrelation models: Some Monte Carlo results. Sociological Methods & Research, 13(2), 155–200. https://doi.org/10.1177/0049124184013002001
- Ghantoji, S. S., Sail, K., Lairson, D. R., DuPont, H. L., & Garey, K. W. (2010). Economic healthcare costs of infection: A systematic review. Journal of Hospital Infection, 74(4), 309–318. http://dx.doi.org/10.1016/j.jhin.2009.10.016.

#### References

- Luke, D. A. & Harris, J. K. (2017). Network Analysis in Public Health: History, Methods, and Applications. Annual Review of Public Health, 28, 69-93. https://doi.org/10.1146/annurev.publhealth.28.021406.144132
- Mayo Clinic Staff. (2016, June 18). C. difficile infection. Retrived from http://www.mayoclinic.org/diseases-conditions/c-difficile/home/ovc-2020226 4
- Simmering, J. E., Polgreen, L. A., Campbell, D. R., Cavanaugh, J. E., & Polgreen, P. M. (2015). Hospital transfer network structure as a risk factor for Clostridium difficile infection. Infection Control and Hospital Epidemiology, 36(9), 1031-1037. http://doi.org/10.1017/ice.2015.130

### Acknowledgements

- Dr. Daniel Sewell
- Dr. Gideon K. D. Zamba
- Dr. Jacob Simmering
- Miles Dietz-Castel & Terry Kirk
- Monica Ahrens
- The University of Iowa Biostatistics Department
- ISIB Program sponsored by the National Heart Lung and Blood Institute Grant: HL131467







National Heart, Lung, and Blood Institute

#### Normal Q-Q Plot



Theoretical Quantiles