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
  ○ Nausea
  ○ Fever
  ○ Dehydration (can lead to kidney failure)
  ○ Death

● How it spreads
  ○ Spores passed in feces when infected don’t wash hands; spread to surfaces
  ○ Touching contaminated surfaces and then touching mouth, nose, or eyes
What is C. diff?

- **Risk factors**
  - Advanced age
  - Prior antimicrobial use
  - Severity of illness
  - Length of stay
  - Being in a hospital with more patients
  - Being in a network of hospitals with more patients

- **How it is treated**
  - Antibiotics
  - Surgery
  - 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**
  - 410 California hospitals identified by AHAID
  - Monthly counts (2008-2011) for number of primary and secondary C. diff cases
  - Average monthly length of stay
  - Monthly percent of patients over 65
  - Other demographic variables

- **Hospital data - locations**
  - Data from hospitals nationwide
  - AHAID and latitude/longitude of each hospital

- **Transfer list**
  - 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
  - CDI mean monthly ranged from 0 to 72 cases
  - Mean monthly length of stay ranged from 1.3 to 76.5 days
  - Mean monthly % over 65 ranged from 0 to 93 percent
  - Mean monthly admissions ranged from 3 to 4619 days
Model - Linear Network Autocorrelation Model

\[ y_i = X_i'^{\beta} + \rho \sum_{j \neq i} A_{ij} y_j + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2) \]

\( y_i \) 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 \( i \)

\( \beta \) is a vector of covariate regression coefficients

\( \rho \) 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

\[ y = X\beta + \rho Ay + \epsilon, \quad \epsilon \sim N(0, \sigma^2) \]

Multivariate normal distribution:

\[ y \sim N((I - \rho A)^{-1}X\beta, \sigma^2(I - \rho A)^{-1}(I - \rho A')^{-1}) \]
Model - Add in Spatial Component

- This is the statistical model that we used:

\[ y = X\beta + \rho A y + \epsilon, \quad \epsilon \sim N(0, \sigma^2 \Sigma_\phi) \]

\[ \Sigma_\phi \] \[ i,j \] = \[ e^{-\phi d_{ij}} \]

\[ y \sim N((I - \rho A)^{-1} X\beta, \sigma^2 (I - \rho A)^{-1} \Sigma_\phi (I - \rho A')^{-1}) \]

- \( d_{ij} \) - distance between hospitals using longitude & latitude
- \( \phi \) - spatial parameter
## Results - 95% CI for Model Coefficients

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intercept</strong></td>
<td>-4.35</td>
<td>-5.39</td>
<td>-3.31</td>
</tr>
<tr>
<td><strong>Mean Length of Stay</strong></td>
<td>0.03</td>
<td>-0.02</td>
<td>0.08</td>
</tr>
<tr>
<td><strong>Mean % Over 65</strong></td>
<td>9.14</td>
<td>7.0</td>
<td>11.28</td>
</tr>
<tr>
<td><strong>Mean Admissions</strong></td>
<td>0.009</td>
<td>0.008</td>
<td>0.01</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.00007</td>
<td>0.00004</td>
<td>.0001</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>184.9</td>
<td>137.0</td>
<td>244.7</td>
</tr>
<tr>
<td>$\log$(Spatial $\phi$)</td>
<td>5.0</td>
<td>-24.81</td>
<td>34.82</td>
</tr>
</tbody>
</table>
Results - Counterfactual Research Question

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

\[
\frac{E(y'1|\rho = 0) - E(y'1|\rho \neq 0)}{E(y'1|\rho \neq 0)} = \frac{(X\beta)'1 - ((I - \rho A)^{-1}X\beta)'1}{((I - \rho A)^{-1}X\beta)'1} = -0.089 \rightarrow -9%
\]

\sim 3,000 \text{ fewer cases per year}

\sim \$16.5 \text{ 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

References


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