CDI Case Detection using Machine Learning and Agent Based Network Modeling

Vivian Ma

Dr. Anil Vullikanti and Dr. Methun Kamruzzaman

Background

Hospital Acquired Infections (HAI's)

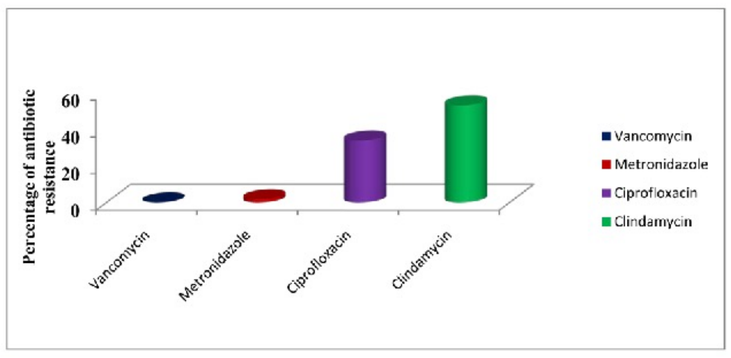
- Infections that develop during a patient's stay at the hospital due to favorable hospital environments
- Common types are C. difficile, pneumonia, bloodstream infections, surgical infections, and urinary tract infections

C. Difficile Infections (CDI)

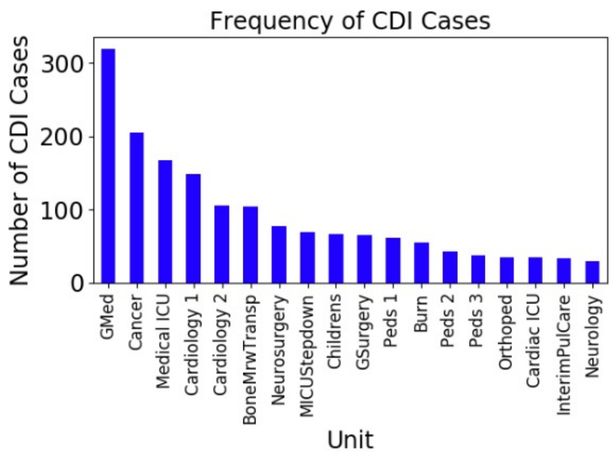
- Can be recovered from the environment surrounding patients with CDI and from healthcare workers who are caring for patients with CDI
- C. difficile spores are resistant to traditional cleaning agents and can stay in the environment for long periods of time
- Colonization pressure is the measure of the proportion of patients infected in a particular area for a specified period of time

Project Goal

- Develop models to predict the infection rates, patient risk to HAIs, and evaluate ways to control the spread of these infections
- Specifically focusing on C. difficile infections (CDI)
- We hypothesize that C. difficile pressure decreases as distance increases



Percentage of antibiotic resistance among different drugs against C. difficile



Experiment Set Up

Set Up

- We wanted to track the total amount of patients infected with C. difficile a patient who stayed at the hospital will encounter during their stay at the hospital and what factors would affect it
- Using data from the UVA hospital, I created a dataset on patient information including all the features listed below
- Each patient has a unique ID that was used to combine all the data on them

Issues Encountered

- Many observations didn't have data for a few of the features means I couldn't include them in the dataset which shortened the dataset a lot
- The model is more likely to be biased because of the smaller sample size

Feature	Definition
PAT_ENC_CSN_ID	Unique patient ID
Age	
Gender (M, F)	
Admission Type (Elective, Emergency, Urgent, Routine)	How they were admitted to the hospital
Total Infected Encountered	Total number of patients with CDI that were encountered during their stay
Charlson Comorbidity Index (0, 1-2, 3-4, >=5)	Method of categorizing comorbidities of patients with six representing the most severe morbidity.
Gastric Acid Suppressors (PPI, HRA)	Whether the patient has them
Hospital admission time and length of stay	When the patient arrived and how long they stayed
CDI Test (ORD_VAL)	Whether they tested positive or negative
Admission/Discharge Source (To_HCF, From_HCF, Internal_Shift)	Where the patient came from and where they are discharged to

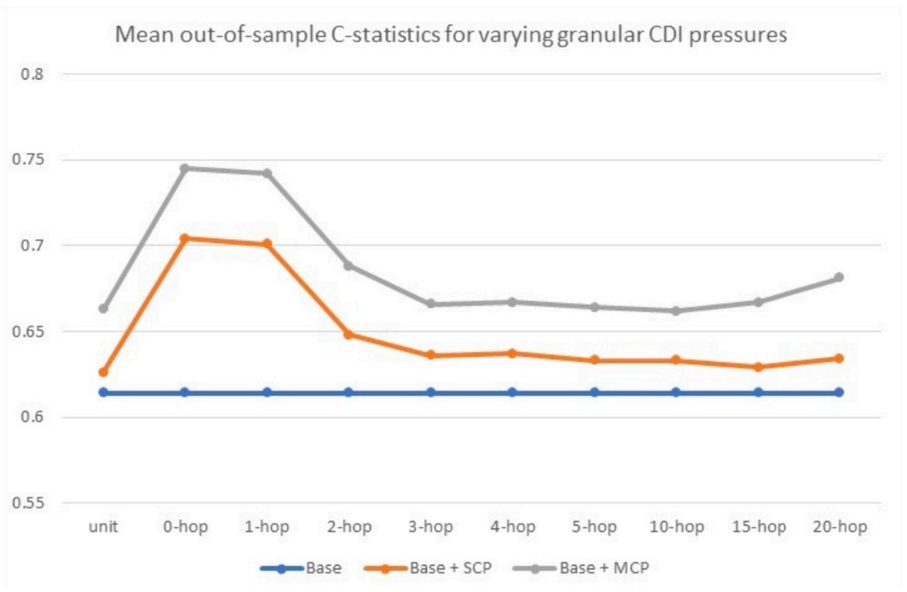
Model and Results

Model

- Used ensemble method to divide the majority class (negative test) into subsets of equal size to minority class (positive test)
- Used k-fold cross validation to train multiple models and test them on each group
- Currently still finishing the models and testing

Expected Outcomes

- CDI pressure mainly arises from roommates and adjacent rooms
- As distance grows the level of association between CDI pressure and CDI acquisition declines
- Different units will have varying numbers of CDI cases



Future Work

- Improving the model by adding features that influence the spread of CDI and removing features that don't influence the spread
- Collecting more data observations to decrease the population bias
- Focusing on different HAIs that are common among patients

References

- Riaz, Talal, et al. "Highly Local CDI Pressures as Risk Factors for CDI." 11 Apr. 2021.