Antibiogram Resistance Pattern Detection And Tracking

Undergraduates: Saarthak Gupta, Caitlyn Fay
Mentors: Dr. Anil Vullikanti, Dr. Chen Chen
Biocomplexity Institute

Background

- **The antibiogram test results provide the drug resistance status of the patient. It can be represented as a vector with the following resistance states:**
  - **NULL–the resistance is unknown**
  - **Sensitive (S)–the patient is not resistant to the drug**
  - **Intermediate (I)–the patient is in the intermediate state, but generally not resistant to the drug**
  - **Resistant (R)–the patient is resistant to the drug**

  Figure 1 shows a sample antibiogram test result.

Drug	Amoxicillin	Clindamycin	Erythromycin	...	Vancomycin
State	NULL	I	R	...	S

Figure #1

- **To effectively track the spread of the pathogen it is critical to identify significant spatio-temporal patterns in antibiogram resistance.**

Results

- **Significant rules found so far include:**

Ampicillin/sulbactam -> Oxacillin	**Ciprofloxacin & Penicillin -> Oxacillin**
Levofloxacin -> Ciprofloxacin	**Ciprofloxacin & Penicillin -> Cefazolin**
Ampicillin -> Penicillin	**Clindamycin -> Erythromycin**
Cefazolin -> Oxacillin	**Erythromycin & Penicillin -> Oxacillin**
Ceftriaxone -> Amoxicillin/clavulanate	**Clindamycin -> Levofloxacin**

- **Example: spatial distribution of rule based on HSA city (figure 5).**

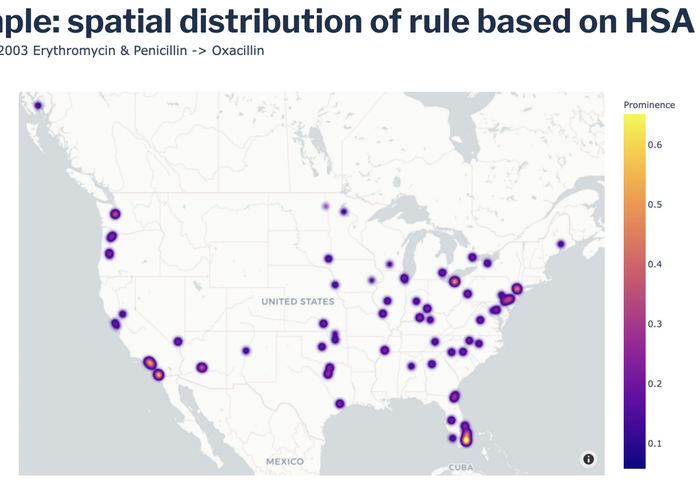


Figure #5

- **Studied temporal distribution of mined rules (figure 6).**

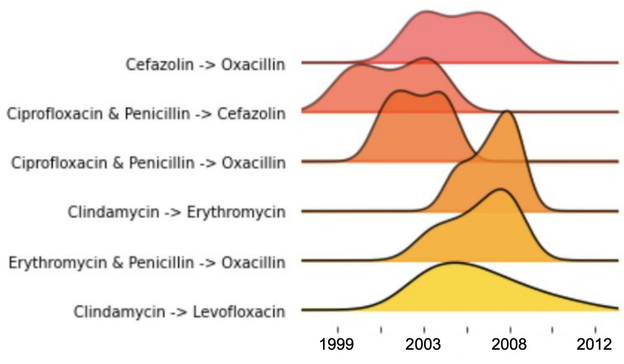
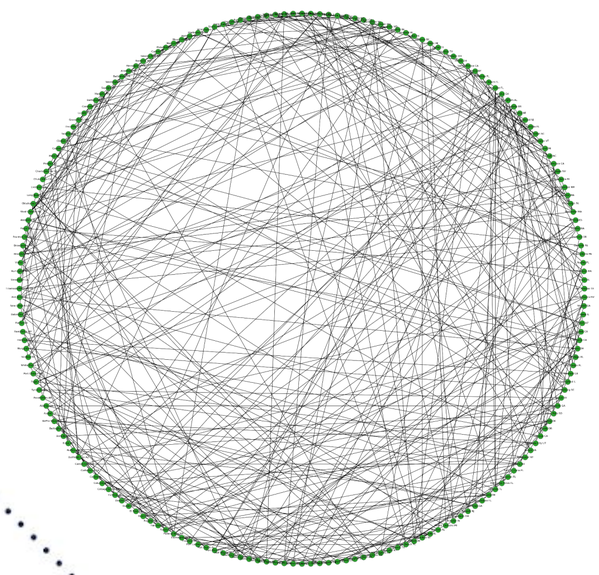


Figure #6

- **Using the haversine formula, the shortest distances between HSA cities were found based on the geocoded data.**
- **A spatial network of HSA cities was constructed (figure 7), treating each HSA city as a node and an edge exists between two cities that are geologically adjacent.**
- **Studied significant pattern flow on the spatial network.**



Section of Network, Enlarged

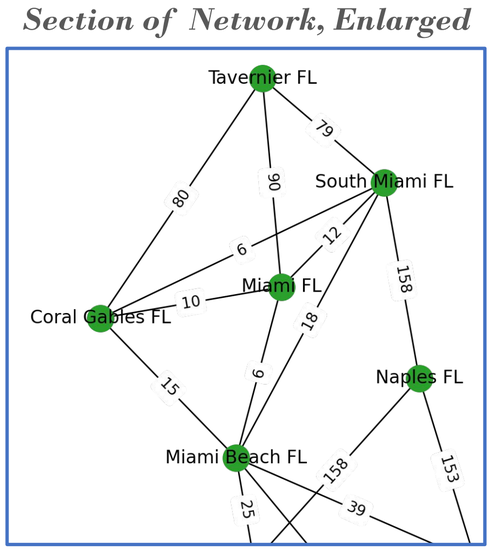


Figure #7

Project Goals

- **Run association data mining tools on the data to find dependencies between antibiotic drugs.**
- **Track the spread of these dependencies over time and space.**
- **Use these results to find methods of predicting the spread of antibiotic resistance in the future.**

Current Work

- **Dataset contains 2.5 million antibiogram test results across 14 years (1999-2012) for 22 drugs tested on Staph Aureus.**
- **Spatially partitioned raw data into HSA cities based on Dartmouth Atlas Data using zip codes.**
- **Some cities are missing data for certain years and each year has some cities with < 100 records (figure 2).**

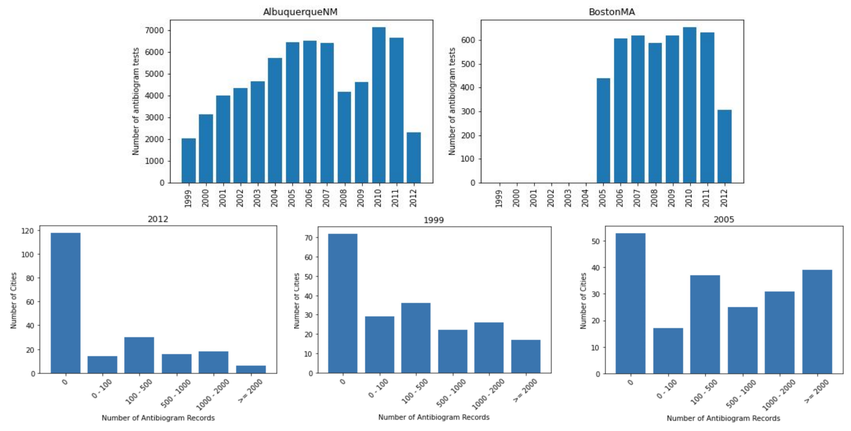


Figure #2

- **Association Rule Mining aims to observe frequently occurring patterns, correlations, or associations from datasets.**
- **Describes dependency of one data item on another data item: A -> B. The dependency can be thought of as an antecedent (If) and a consequence (Then) statement.**
- **Mapped each drug to a number for association rule mining. If a particular test result showed up as "R" it was included as true in the input vector, false otherwise.**
- **Performed yearly association rule mining for each HSA city. Sample association rules are displayed in figure 4.**
- **Used a geocoding API (Nominatim) to find OSM data for each city. Regions are unevenly spread but represent the population density (figure 3).**
- **Studied spatial and temporal patterns in the mined rules and converted significant rules back to drugs using the mapping. Explored clinical significance of the observed patterns.**

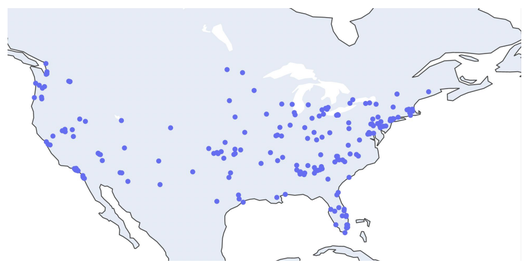


Figure #3

Figure #4

Future Work

- **Split data into groups depending on Oxacillin resistance in order to determine how antibiogram patterns differ with MRSA / MSSA strains (hospital/community).**
- **Find significant antibiogram patterns and identify specific patterns for future use in prediction models.**
- **Write a research paper of findings alongside Eili Klein, clinician at Johns Hopkins.**

References

1. Wilhelmiina Hämäläinen. Kingfisher: an efficient algorithm for searching for both positive and negative dependency rules with statistical significance measures. Knowledge and information systems, 32(2):383–414, 2012.
2. Dartmouth Atlas Project. Dartmouth atlas data. https://data.dartmouthatlas.org/supplemental/#crosswalks.
3. Wilhelmiina Hämäläinen. Kingfisher - a tool for searching statistical dependencies/association rules. https://sites.google.com/site/whsivut/ home?authuser=0.