

Spatiotemporal Antibiotic Resistance Pattern Monitoring using Geographical Information System based Hierarchical Cluster Analysis

Roshan Hewapathirana^a, Gamini Wijayarathna^b

^a Post Graduate Institute of Medicine, University of Colombo, Sri Lanka

^b Department of Industrial Management, Faculty of Science, University of Kelaniya, Sri Lanka

Abstract

Bacterial antimicrobial resistance in both the medical and agricultural fields has become a serious problem worldwide. Antibiotic resistant strains of bacteria are an increasing threat to human health, with resistance mechanisms having been described to all known antimicrobials currently available for clinical use. Monitoring the geotemporal variations of antibiotic resistance pattern is crucial factor in planning a successful therapeutic guidelines preventing further emergence of antibiotic resistance. This study is based on the retrospective spatiotemporal analysis of laboratory results of Antibiotic Sensitivity Tests, time stamped with the date and time of the microbiological specimen dispatched to the laboratory. Geographic location of the isolated bacterial colony is specified with the latitude and the longitude of the patient's location. Agglomerative Hierarchical Clustering was performed on antimicrobial resistance findings based on the geographic locations generating series of Heatmaps to visualize the extent of the resistance pattern. Sequential Hierarchical cluster analysis was proven to be effective in visualization of antibiotic resistance using Heatmaps demonstrating the temporal variations of the antibiotic resistance patterns.

Keywords:

Antibiotic sensitivity test, Antibiotic resistance, Spatiotemporal pattern, Cluster analysis, Agglomerative hierarchical clustering, Geographic information system, Heatmap.

Introduction

The emergence of antimicrobial resistance is neither an unexpected nor a new phenomenon. It is, however, an increasingly problematic situation due to the frequency with which new emerging resistance phenotypes are occurring among many bacterial pathogens.

Over the past 6 decades, the introduction of new classes or modifications of older classes of antimicrobials has been matched slowly but surely by the development of new bacterial resistance mechanisms, [1].

Antibiotic Sensitivity Test is the investigation carried out to determine which antibiotic will be most successful in treating a bacterial infection in vivo. Testing for antibiotic sensitivity

is often done by the Kirby-Bauer method, [2] where small wafers containing antibiotics are placed onto a plate upon which bacteria are growing. If the bacteria are sensitive to the antibiotic, a clear ring, or zone of inhibition, is seen around the wafer indicating a poor growth.

Numerous retrospective and prospective studies have demonstrated that increases in antimicrobial resistance occur among both pathogenic and commensal bacteria after introduction of an antimicrobial agent, [3]. Infections caused by resistant bacteria have been shown to be more frequently associated with increased morbidity and mortality than those caused by susceptible pathogens. The emergence and dissemination of bacterial antimicrobial resistance is the result of numerous complex interactions among antimicrobials, microorganisms, and the surrounding environment including geographical location and population dynamics.

There are considerable variations in the antibiotic susceptibility of clinically important bacterial strains based on the geographical distribution of hosts, [4, 5] and exposure to environmental factors over a period of time, [6].

Geographic information systems (GIS) and analysis based on GIS has become widespread and well accepted tool in the field of medicine as well. Geographic information systems are an important way in which to better illuminate how humans interact with their environment to create or deter health, [7]. Recent advancements in the analysis of disease maps have been influenced by and benefited from the adoption of new practices for georeferencing health data and new ways of linking such data geographically to potential sources of environmental exposures, the locations of health resources and the geodemographic characteristics of populations, [8].

It was shown in relation to some bacterial strains that the geographic maps of antibiotic resistance can be used to guide physician antibiotic selection before culture results are available. This has significant implications for the health care provider in proper antibiotic selection within the community minimizing the treatment failures and emergence of new, drug resistant bacterial strains, [9].

Methods

Antibiotic Sensitivity Test results extracted from a retrospective laboratory based study was the key data set concerned in this analysis. When an Antibiotic Sensitivity Test is performed, type of the specimen, the name of the bacterial species isolated and the resistance or the sensitivity of the bacteria towards a set of selected antimicrobial agents along with the personal information of the patient from whom the specimen was collected are recorded in all microbiological laboratories. Number of antibiotics a bacterial colony is tested for ranges from 8 to 12 depending on the laboratory where the Antibiotic Sensitivity Test was performed. The types of antimicrobial agents used for the Antibiotic Sensitivity Tests are also laboratory dependent unless a request is made by the physician to include a specific antibiotic based on the clinical condition and suspected pathogen.

Out of the above mentioned information available in microbiological laboratories, type of the specimen, date and time when the specimen was dispatched to the laboratory, the bacterial species isolated, resistance or sensitivity to each antimicrobial agents and the location of the patient were fed in to a MySQL based relational database using web based interface developed with scripting languages *PHP*: Hypertext Preprocessor, JavaScript and Python. Location of the patient for each test result was linked to a look up table in order to match location from where the microbiological sample was collected to its nearest possible latitude and longitude.

The interface developed has the capacity to filter the antibiotic sensitivity and resistance counts for each spatial location based on latitude and longitude and the type of antimicrobial agent. Several pre-formatted queries are provided through the interface to find counts based on bacterial strains, antimicrobial agent and time period.

Google Maps, [10] was used as the primary GIS layer as shown in Figure 1. Google Maps provides a Java Script based API which includes calls to add markers on the map for various data points.

However a supportive API, Gheat, [11] has to be employed to assist Google Maps layer, since the study manipulates vast number of data points and mapping those antibiotic resistance counts is beyond the capacity of Google Maps alone. An additional tile set was created with the aid of Gheat and which was layered on top of the base map tile set to visualize clustered antibiotic resistance counts of each antimicrobial agent based on the latitude and longitude as a Heatmap.

Similar to the Google Maps API, Gheat creates the final tile layer using multiple 256x256 pixel tiles knitting them together. Since the data points are time stamped, it is possible to generate these tile sets based on the time stamp of the individual Antibiotic Sensitivity Test result in a serial manner to analyze the periodical distribution of the resistance counts.

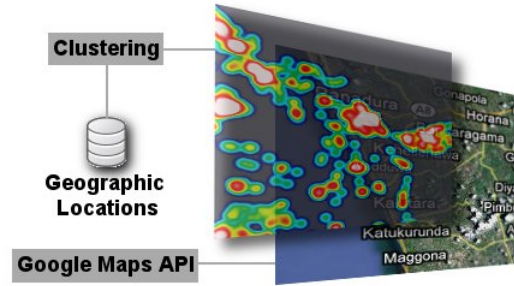


Figure 1 – Use of Google Maps API and Clustering geographic locations of test results

While frequency of occurrence are being calculated for each geographical point, those locations have to be grouped to different hierarchical layers based on the frequency of the occurrence of resistance or sensitivity reported for each cluster. In order to do this the interface accepts the individual occurrence of resistance from the database and cluster the data points, based on Agglomerative Hierarchical Clustering, [12] algorithm. The frequency clustering is possible as a cross sectional study with a single observation or as a longitudinal study with sequence of observations. In clustering antibiotic resistance results, Euclidean distance, between two data points is considered the preferred Metric as in Equation (1) to decide dissimilarity between sets of observations.

$$\|a - b\|_2 = \sqrt{\sum_i (a_i - b_i)^2} \quad (1)$$

The selected Linkage Criteria to decide the distance between sets of observations is Single Linkage (Nearest Neighborhood) where pair wise distance between observations is used as shown in Equation (2).

$$\min\{d(x, y) : x \in A, y \in B\} \quad (2)$$

The distance between two clusters is calculated as the distance between the two closest elements in each cluster. Linkage function – the distance $d(x, y)$ between cluster A and cluster B is described by, $d(A, B) = \min(d(x, y))$ where,

- $d(x, y)$ is the distance between element $x \in A$ and $y \in B$;
- A and B are two sets of clusters within the same hierarchy

The maximum positional accuracy of storing the values for the latitude and longitude values are to the $1/1000000^{\text{th}}$ of a degree. So 0.000001 degrees of latitude or the longitude would be the closest distance for any two elements to be identified as two different clusters, as specified by the $d(x, y)$ in Equation (2) above.

Gheat color scheme consists of four channels of colors (red, green, blue, and alpha). Each channel ranges from 0 to 256. When the data is being clustered alpha channel is used to draw the base template of the Heatmap as shown in Figure 2. Then

each cluster will be colorized using appropriate combination of red, green and blue channels according to the color scheme. Once the color scheme is applied, each cluster of the Heatmap will have a unique color combination as shown in Figure 2.

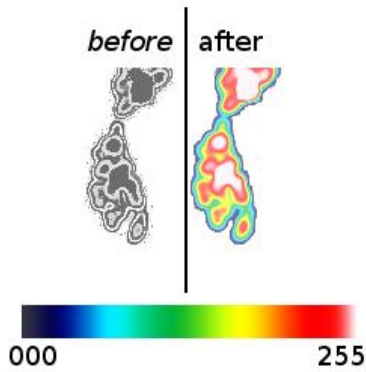


Figure 2 – Before and after colorizing clusters creating a Heatmap

This approach allows agglomerative positioning of clusters where each observation starts in its own cluster and pairs of clusters are merged as the process moves up the hierarchy. In these hierarchies of clusters, least dense clusters would occupy the bottom layers where as high density clusters would occupy the top layers. A color scale can be defined shown in Figure 2 to colorize the each layer in the Heatmap.

Results

Heatmap based on clustered antibiotic resistance counts for each bacteria. Cases were mapped with reference to the latitude and longitude of the specimens collected as shown in Figure 3.

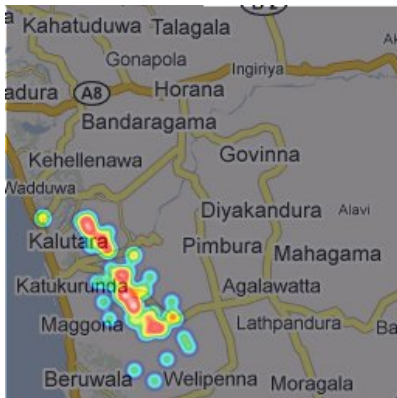


Figure 3 – Antibiotic resistance cases mapped to the respective GIS positions

Data extracted from the Antibiotic Sensitivity Test result was fed in to the GIS system and the interface was able to generate

Similarly, cases resistant to a selected antimicrobial agent can be filtered by time stamp and Heatmap can be generated sequentially. In this approach spatial distribution of the antibiotic resistance cases can be visualized periodically as shown in Figure 4.

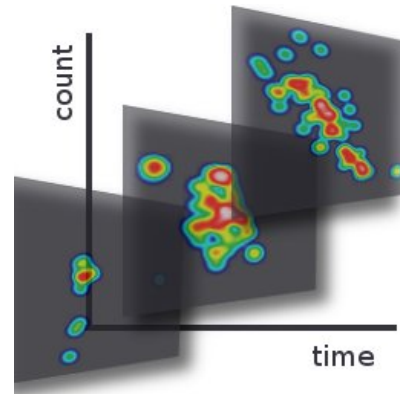


Figure 4 – Sequential Heatmap highlighting the changes of the resistance counts against time

Discussion

The serial Heatmaps generated by the interface are useful in cross sectional studies of the database, providing visual clues on the extent of the distribution of antibiotic resistant cases at any given time.

With the use of Heatmaps, large number of values can be visually presented compares to conventional map markers used with the Google Map API.

There are numbers of clinically significant statistical analysis can be performed with the help of pre-formatted database queries.

- Comparison of resistance or sensitivity of a bacteria to an antibiotic in different points in time as a longitudinal study
- Comparison of resistance or sensitivity of a antibiotic to one or more bacteria as a cross sectional study
- Comparison of resistance or sensitivity of a antibiotic to one or more bacteria as a longitudinal study
- Comparison of resistance or sensitivity of a bacteria to one or more antibiotic as a cross sectional study
- Comparison of resistance or sensitivity of a bacteria to one or more antibiotic as a longitudinal study

Conclusion

The study was able to demonstrate the ability to incorporate Geographical Information Systems with clinical and laboratory based data. Cross sectional as well as longitudinal analysis is possible with the use of cluster mapping.

Heatmaps are easy to interpret than numerical data or charts in representing pattern of distribution of antibiotic resistance cases over a geographical location.

It was also shown that, in longitudinal sampling of frequencies, a temporal analysis of the antibiotic resistance or sensitivity pattern is possible with sequential Heatmaps.

Acknowledgments

Authors would like to make this an opportunity to thank Chad W. L. Whitacre for the support rendered on Gheat API and staff of all microbiological laboratories for providing the access to Antibiotic Sensitivity Test Results.

References

- [1] White DG, Cray FP, and Chiller CT. The National Antimicrobial Monitoring System (NRMS). NMC Annual Meeting Proceedings 2006: 56 – 60.
- [2] Cappuchino JG, Microbiology, A Laboratory Manual. 7th ed. Pearson Education, 2007.
- [3] Hawkey PM and Jones AM. The changing epidemiology of resistance. *Journal of Antimicrobial Chemotherapy* 2009; 64: 13 – 10.
- [4] Sannes MR, Kuskowski MA and Johnson JR. Geographical distribution of antimicrobial resistance among *Escherichia coli* causing acute uncomplicated pyelonephritis in the United States. *FEMS Immunol Med Microbiol* 2004;42(2):213-8.
- [5] García-Rey C, Aguilar L, Baquero F, Casal J and Dal-Ré R. Importance of local variations in antibiotic consumption and geographical differences of erythromycin and penicillin resistance in *Streptococcus pneumoniae*. *J Clin Microbiol* 2002; 40(1):159-64.
- [6] Chambers HF and Deleo FR. Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nat Rev Microbiol* 2009; 7(9):629 – 641.
- [7] Ricketts TC. Geographic information systems and public health. *Annu Rev Public Health* 2003; 24:1 - 6.
- [8] Rushton G. Public health, GIS, and spatial analytic tools. *Annu Rev Public Health* 2003;24:43 – 56.
- [9] Tirabassis MW, Wadle G, Moriarty KP, Garb J, Konefal RA, Sachs BF and Walt R. Geographic Information System localization of community-acquired MRSA soft tissue abscesses. *Journal of Pediatric Surgery* 2005; 40: 962 – 966.
- [10] Google Maps API [Online] 2009. [cited 2009 February 10]; Available from: URL: <http://code.google.com/apis/maps/>
- [11] Gheat, heatmaps for Google Maps [Online]. 2008 April 29 [cited 2009 March 2]; Available from: URL: <http://code.google.com/p/gheat/>
- [12] Han J and Kamber M. *Data mining, Concepts and Techniques*. 2nd ed. Morgan Kaufman Publishers, 2006.

Address for correspondence

Roshan Hewapathirana
 Post Graduate Institute of Medicine
 No. 160
 Norris Canal Road,
 Colombo 07, Sri Lanka
roshanhewapathirana@gmail.com