Visualizing and Monitoring Data in BioSense Using SaTScan

Colleen A. Bradley, MSPH[1,2], BioIntelligence Center Data Monitor Group[1]*

Centers for Disease Control and Prevention[1], SAIC[2]

Objective

The objective is to describe the visualization and monitoring of the national spatio-temporal SaTScan results in the BioSense application. This is the first application of this algorithm to a national early event detection and situational awareness system.

Background

BioSense data includes Department of Defense and Veterans Affairs ambulatory care diagnoses and procedures, as well as Laboratory Corporation of America lab test orders. Data are mapped to eleven syndrome categories [1]. SaTScan is a spatio-temporal technique that has previously been applied to surveillance at the metropolitan area level [2,3]. Visualization of national results involves unique issues, including displaying cluster information that crosses jurisdictions, zip codes with highly variant data volume, and evaluating large multiple state clusters. SaTScan was first implemented in June 2005 in the BioSense application for daily monitoring at CDC's BioIntelligence Center (BIC).

Methods

BioSense applies the SaTScan algorithm to each national data source and displays results for all states and metropolitan areas (MRA) in jurisdictionally based views. The nation is divided into distinct geographical units ("the grid") based on population density. Small Area Regression and Testing (SMART) modeling is the basis by which expected counts are generated for each grid [4]. A maximum of ten clusters are displayed for each source and syndrome. Clusters can incorporate up to 7 days worth of data. BIC monitors provided input and feedback regarding the data visualization components and results to increase utility for surveillance. Methods for visualizing cross-jurisdictional clusters were developed.

Results

BioSense indicates the number of SaTScan clusters by data source and syndrome in a table on the home page. Users may navigate to a syndrome-specific cluster page, which includes a jurisdictional cluster map and a table for each data source that lists summary information for all clusters. The cluster circle on the map indicates if a cluster extends beyond a state or MRA; however, detailed record information for patients outside that jurisdiction is not displayed.

A user may select a cluster to obtain a detailed data table, including the ratios between the observed and expected counts for each day in the cluster. Daily cluster record counts are linked to the detailed line listing, including diagnostic, demographic, facility, and patient information.

Conclusions

BIC monitoring of BioSense SaTScan results indicates several issues. Cluster sizes can be geographically large, especially in the less populated mid-Western US. Evaluating clusters for clinical and/or epidemiologic relevance can be difficult and requires monitor ability to: 1) view the detailed line list for all records in the cluster, 2) graph historical data for zip codes in the cluster for longer term temporal analysis, and 3) communicate with state and local public health officials. In determining the number of clusters to display, there must be a balance between monitoring time and potential cluster importance. Monitor feedback will continue to be utilized to address large cluster size, improve data visualizations, and enhance end user analysis capabilities. Upon further development, SaTScan results and jurisdictional visualizations will be released for state and local public health monitoring in the BioSense application, and user feedback will be solicited regarding enhancements.

References

[1] Syndrome Definitions for Diseases Associated with Critical Bioterrorism-Associated Agents. Accessed at: http://www.bt.cdc.gov/surveillance/syndromedef/index.asp

[2] Kulldorff M. Prospective time periodic geographical disease surveillance using a scan statistic. J R Statis Soc A 2001;164:61-72.

[3] Mostashari F, Kulldorff M, Hartman J, Miller J, and Kulasekera V. Dead Bird Clusters as an Early Warning System for West Nile Virus Activity. Emerg Infect Dis 2003:9(6):641-646.

[4] Kleinman K, Lazarus R, Platt R. A generalized linear mixed models approach for detecting incident clusters of disease in small areas, with an application to biological terrorism. Am J Epidemiol 2004;159:217-24.

Further Information:
Colleen A. Bradley, cbradley1@cdc.gov

* BioIntelligence Center Monitor Group
Leslie Sokolow, MSc, CGIS
Shirley Willson, MIT
Martin Fayomi, MPH
Yukun Jin, MS
Matthew Miller, BA, EMT-P
Duane Zomer, BS