

Algorithms to Characterize Syndromic Surveillance Spatial Alerts

Jingsong Lu M.S., Kevin Konty M.S, Chris Goranson B.A., Farzad Mostashari M.D. M.S.c

Bureau of Epidemiology Services, New York City Department of Health and Mental Hygiene

OBJECTIVE

This paper explores some visualization methods for characterizing spatial signals detected by SaTScan and discusses how these maps might aid in deciding whether to investigate a signal, as well as the scope and focus of the investigation.

BACKGROUND

Syndromic surveillance systems generate temporal and spatial alerts when statistical thresholds are exceeded. Although it is important to rapidly identify true public health emergencies and respond to them appropriately, fully investigating all signals would be extremely time-consuming and inefficient. A number of approaches have been developed by New York City (NYC) ED surveillance to characterize spatial signals prior to initiating an investigation. These include: examining the distributions of cases within the cluster by sub-syndrome, hospital, sex, and age; comparing spatial signals of related syndromes or for consecutive days of the same syndrome; and reviewing line lists. Maps are occasionally produced to investigate particularly worrisome signals. These maps might be more useful in less severe cases where the necessity of an investigation is unclear. Plus, since the format of the maps is constant, they can be automated for use in a wider spectrum of cases. We automate the production of maps displaying raw data, signal details, and demographic information.

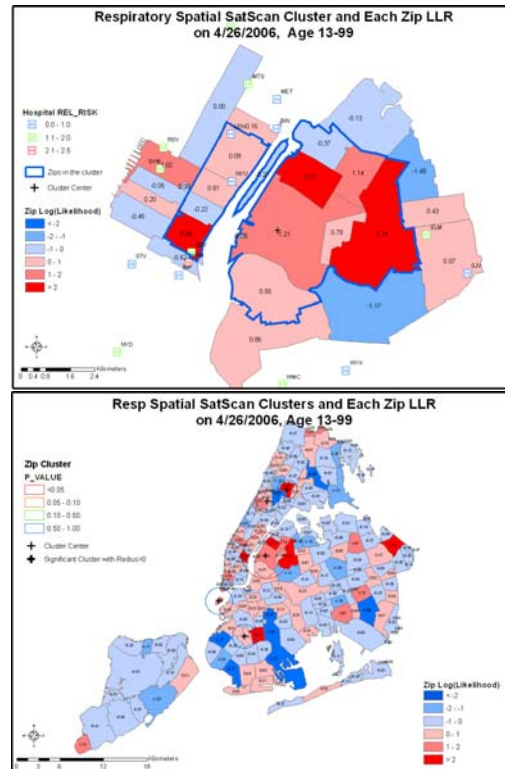
METHODS

NYC ED surveillance receives data from 48 hospitals encompassing approximately 90% of all NYC ED visits. Syndromes are coded according to chief complaint and spatial (zip, hospital) analyses are performed using SaTScan. For this project, python scripts run within SAS programs to create layers for an output GIS. These layers include raw and expected values, SaTScan output, population data, intermediate SaTScan values such as zip and hospital log-likelihood ratios computed in SAS. We also quantify the contribution of individual geographic areas to a cluster's significance. The same layers are created by age and for related syndromes. Time series maps (or video) are used to monitor the outbreak.

RESULTS

The resulting maps allow for the examination of signals details alongside large amounts of contextual information. The "naked eye" view of the data becomes more valuable when presented with the cluster and its details, in turn providing a better picture of the signal. For example, the figures below

present a respiratory zip level signal in the 13+ year olds. The cluster identified can be seen to be highly heterogeneous -- only 5 of 37 constituent zip codes have LLRs greater than one. The Veterans Affairs Manhattan (VAM) near the zip with the highest LLR is the only hospital that has a RR greater than two.



The age-group analysis identifies a similar cluster in the 40-64 year olds involving the same 5 zip codes, while no clusters were detected among adults in other age groups. We may focus on those zip codes and call VAM if we decide to start an investigation. Other layers provide raw and expected values for all zip codes, intermediate SaTScan output, and previous signals for the same syndrome.

CONCLUSIONS

The programs automate mapping procedures previously only performed for severe outbreaks. Using python scripts running within SAS allows for more control of spatial tools, ensures extensibility, and improves the timeliness of signal characterization and decision-making for response. The resulting GIS provides instant visualization of SaTScan signal detail and intermediate calculations. This results in a highly useful decision support system.

Further information: Jingsong Lu, jl@health.nyc.gov