Cluster Detection Comparison in Syndromic Surveillance

C. Goranson[1,3], Takahashi K.[2], Tango T. [2], A. Cajigal[1], M. Paladini[1], E. L. Murray[1], T. Nguyen[1], K. Konty[1], F. Hardisty[3]

New York City Department of Health and Mental Hygiene(NYCDOHMH), NY, USA[1] , National Institute of Public Health, Wako-shi, Saitama, Japan[2], The Pennsylvania State University[3], PA, USA

OBJECTIVE

To use the New York City Department of Health and Mental Hygiene's (NYC DOHMH) emergency department (ED) syndromic surveillance data to evaluate FleXScan's flexible scan statistic and compare it to results from the SaTScan circular scan. A second objective is to improve cluster detection in by improving geographic characteristics of the input files.

BACKGROUND

The NYC DOHMH collects data daily from 50 of 61 (82%) EDs in NYC representing 94% of all ED visits (avg daily visits ~10,000). The information collected includes the date and time of visit, age, sex, home zip code and chief complaint of each patient. Observations are assigned to syndromes based on the chief complaint field and are analyzed using SaTScan to identify statistically significant clusters of syndromes at the zip code and hospital level [1]. SaTScan employs a circular spatial scan statistic and clusters that are not circular in nature may be more difficult to detect. FlexScan employs a flexible scan statistic using an adjacency matrix design [2][3].

METHODS

Counts of syndrome visits were aggregated at the zip code level for 2005. FleXScan's flexible scan and SaTScan's circular scan were analyzed by comparing the most likely cluster (primary cluster) identified; the secondary clusters identified; location and area of identified cluster; P-value and relative risk. Both projected and unprojected coordinate systems were used to identify sensitivity in clusters to changes in measurement and coordinate systems. Improving the FleXScan matrix file provided a method for capturing area connectivity where bridges, tunnels, or subway lines existed between them. This was not possible to do in SaTScan. ZIP code area centroids were weighted to reflect the underlying population distribution of the areas. Both FleXScan and SaTScan were run again using the reweighted centroids.

RESULTS

FleXScan and SaTScan both detected similar, overlapping areas in three of the time periods investigated. Non-circular clusters with a high relative risk were detected by FleXScan's flexible scan, but this was not detected by SaTScan (Figure 1). However, known clusters were detected at a more significant p-value by SaTScan than FleXScan (p=0.002 vs. p=0.179). Weighting ZIP code centroids based on population and improving the connectivity matrix changed results; over a one week period p-values increased 50% of the time, decreased 36% of the time, and stayed the same 14% of the time when weighted centroids were employed. The differences were most prominent where unweighted centroids had not been representative of underlying population distributions in the areas.

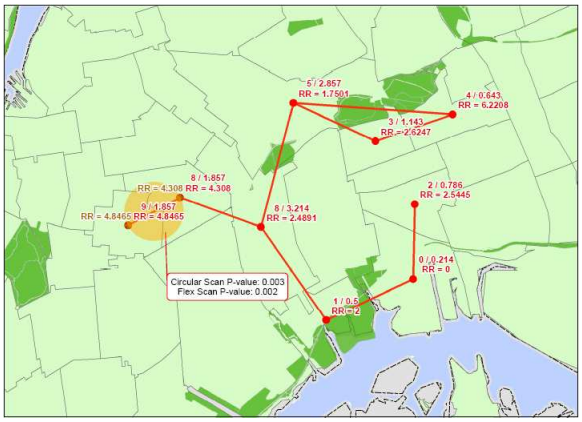

Figure 1 – Circular scan significant cluster (orange) and flexible scan significant cluster (red). Relative risk is also shown.

CONCLUSIONS

Flexible and circular scans are complementary tools that each have their strengths and should be used together. SaTScan provides a useful method for detecting clusters more circular in nature; FleXScan is one approach to better identify non-circular clusters by employing a flexible spatial scan statistic. Used together they may provide the best alternative to characterizing an outbreak.

Updating and improving the quality of area centroid files through population weighting, and the creation of accurate matrix files may improve the accuracy of syndromic surveillance. Additional analyses should be conducted to provide a statistical basis to comparisons between the methods. A second comparison between other confirmed outbreaks will attempt to identify the sensitivity and specificity of each under different circumstances.

REFERENCES

[1] Heffernan R, Mostashari F, Das D, Karpati A, Kulldorff M, Weiss D, Syndromic surveillance in public health practice, New York City. Emerging Infectious Diseases. 2004 May:10(5):858-64.

[2] Tango T, Takahashi K, A flexibly shaped spatial scan statistic for detecting clusters. International Journal of Health Geographics. 2005 May 18;4(1):11.

[3] Takahashi K, Kulldorff M, Tango T and Yih K. A flexibly shaped space-time scan statistic for disease outbreak detection and monitoring. International Journal of Health Geographics 2008, 7:14

This project funded under NSF / EAPSI Award # 0812980.