

Building a Better Syndromic Surveillance System: the New York City Experience

Robert Mathes*, Jessica Sell, Anthony W. Tam, Alison Levin-Rector and Ramona Lall

New York City Department of Health and Mental Hygiene, Queens, NY, USA

Objective

To evaluate temporal and spatial aberration detection methods for implementation in a local syndromic surveillance system.

Introduction

The New York City (NYC) syndromic surveillance system has monitored syndromes from NYC emergency department (ED) visits since 2001, using the temporal and spatial scan statistic in SaTScan for aberration detection. Since our syndromic system was initiated, alternative methods have been proposed for outbreak identification. Our goal was to evaluate methods for outbreak detection and apply the best performing method(s) to our daily analysis of syndromic data.

Methods

We tested six temporal aberration detection methods: a modified C2 algorithm [1], a CUSUM algorithm, the Holt-Winters (HW) smoothing method, a generalized linear model (GLM), a temporal scan statistic in SaTScan [2], and an autoregressive integrated moving average (ARIMA) model. We tested four spatio-temporal methods: a generalized linear mixed model (GLMM), a Bayesian model [3], and the spatial scan statistic using the Bernoulli probability model and space-time permutation scan statistic in SaTScan [2]. We analyzed ED visits that occurred in NYC during 2010-2011. To test a method's ability to detect different outbreak types, we added simulated outbreaks to NYC syndromic data. A total of 180 datasets were created, each containing one spike of varying duration and magnitude, allocated either to a single ZIP code, a cluster of ZIP codes, or citywide. To compare the performance of these methods, we estimated sensitivity, specificity, timeliness, and positive predictive value (PPV) of signals. Timeliness was estimated by calculating the earliest day of detection as a proportion of the outbreak length before the peak number of excess cases (timeliness ranged from 1 to 0, with detection on the first day of an outbreak resulting in a timeliness measure of 1). To determine how difficult it would be to implement these methods into our system, we recorded programming time, computer run time, and skills (detailed coding or statistical) needed to program the method.

Results

The temporal scan statistic, our current method for temporal outbreak detection, performed the best in regard to sensitivity and timeliness of detection (Table 1). Specificity was low however, suggesting the method may signal often, even when there is no outbreak. All methods had poor PPV, so when a method did signal, it was unlikely to be an outbreak. Among the spatial methods, the GLMM had the highest sensitivity (Table 2), though it signaled the most often. Of note, the methods performed better when the outbreaks were of longer duration (>5 days), both in detecting outbreaks at all and the timeliness of outbreak detection. The easiest method to code was the modified C2 and the most difficult was the Bayesian spatial model.

Conclusions

While there was variability in method performance, we feel there is no one best method for outbreak detection, as we have to weigh

issues like signal frequency, specificity, and timeliness. We also found certain methods required time and skill to code, which is useful when deciding which methods to implement into a daily system. Our poor PPV may be explained in that methods are detecting unlabeled outbreaks given we used actual syndromic data. Our next step is to continue comparing results between these methods, and select the most promising methods to run prospectively in our current system.

Table 1. Metrics of the temporal models running at a fixed alert rate of 0.01 (1 alarm/100 days)

	Sensitivity	Specificity	Timeliness	PPV	Computer run time for one day, citywide	Programming time
ARIMA	0.15	0.99	0.27	0.02	11m 8s	3 weeks
GLM	0.32	0.99	0.57	0.06	1m 52s	3 days
HW	0.05	1.0	0.17	0.25	20m 5s	2 weeks
C2	0.32	0.97	0.49	0.02	23s	1 week
CUSUM	0.49	0.84	0.69	0.02	61s	1 week
Temporal scan statistic*	0.78	0.52	0.89	0.01	4s	3 days

*alert generated at $p < 0.01$

Table 2. Metrics of the spatial models

	Sensitivity	Specificity	Timeliness	PPV	Computer run time for one day across all ZIPs	Programming time
Spatial scan statistic	0.08	0.97	0.74	<0.01	6s	3 days
Space-time permutation	0.37	0.64	0.93	<0.01	12s	3 days
Bayesian	0.24	0.86	0.85	0.01	2m	10 weeks
GLMM	0.66	0.52	0.99	0.01	50s	1 week

Keywords

Evaluation; Syndromic surveillance; Aberration detection methodologies

References

1. Tokars, J.L., et al., *Enhancing time-series detection algorithms for automated biosurveillance*. Emerg Infect Dis, 2009. **15**(4): p. 533-9.
2. Kulldorff, M. and I. Information Management Service, *SaTScan v9.1: Software for the spatial and space-time scan statistic* <http://www.satscan.org/>, 2009.
3. Corberan-Vallet, A. and A.B. Lawson, *Conditional predictive inference for online surveillance of spatial disease incidence*. Stat Med, 2011. **30**(26): p. 3095-116.

*Robert Mathes

E-mail: rmathes@health.nyc.gov

