

Category-Specific Comparison of Univariate Alerting Methods for Biosurveillance Decision Support

Yevgeniy Elbert*, Vivian Hung and Howard Burkom

JHUAPL, Laurel, MD, USA

Objective

For a multi-source decision support application, we sought to match univariate alerting algorithms to surveillance data types to optimize detection performance.

Introduction

Temporal alerting algorithms commonly used in syndromic surveillance systems are often adjusted for data features such as cyclic behavior but are subject to overfitting or misspecification errors when applied indiscriminately.

In a project for the Armed Forces Health Surveillance Center to enable multivariate decision support, we obtained 4.5 years of outpatient, prescription and laboratory test records from all US military treatment facilities. A proof-of-concept project phase produced 16 events with multiple evidence corroboration for comparison of alerting algorithms for detection performance.

We used the representative streams from each data source to compare sensitivity of 6 algorithms to injected spikes, and we used all data streams from 16 known events to compare them for detection timeliness.

Methods

The six methods compared were:

- 1) Holt-Winters generalized exponential smoothing method (1)
- 2) automated choice between daily methods, regression and an exponential weighted moving average (2)
- 3) adaptive daily Shewhart-type chart
- 4) adaptive one-sided daily CUSUM
- 5) EWMA applied to 7-day means with a trend correction; and
- 6) 7-day temporal scan statistic

Sensitivity testing: We conducted comparative sensitivity testing for categories of time series with similar scales and seasonal behavior. We added multiples of the standard deviation of each time series as single-day injects in separate algorithm runs. For each candidate method, we then used as a sensitivity measure the proportion of these runs for which the output of each algorithm was below alerting thresholds estimated empirically for each algorithm using simulated data streams. We identified the algorithm(s) whose sensitivity was most consistently high for each data category.

For each syndromic query applied to each data source (outpatient, lab test orders, and prescriptions), 502 authentic time series were derived, one for each reporting treatment facility. Data categories were selected in order to group time series with similar expected algorithm performance:

- 1) Median > 10
- 2) $0 < \text{Median} \leq 10$
- 3) Median = 0
- 4) Lag 7 Autocorrelation Coefficient ≥ 0.2
- 5) Lag 7 Autocorrelation Coefficient < 0.2

Timeliness testing: For the timeliness testing, we avoided artificiality of simulated signals by measuring alerting detection delays in the 16 corroborated outbreaks. The multiple time series from these

events gave a total of 141 time series with outbreak intervals for timeliness testing.

The following measures were computed to quantify timeliness of detection:

1. Median Detection Delay – median number of days to detect the outbreak.
2. Penalized Mean Detection Delay – mean number of days to detect the outbreak with outbreak misses penalized as 1 day plus the maximum detection time.

Results

Based on the injection results, the Holt-Winters algorithm was most sensitive among time series with positive medians. The adaptive CUSUM and the Shewhart methods were most sensitive for data streams with median zero. Table 1 provides timeliness results using the 141 outbreak-associated streams on sparse (Median=0) and non-sparse data categories.

[Insert table #1 here]

The gray shading in the table 1 indicates methods with shortest detection delays for sparse and non-sparse data streams. The Holt-Winters method was again superior for non-sparse data. For data with median=0, the adaptive CUSUM was superior for a daily false alarm probability of 0.01, but the Shewhart method was timelier for more liberal thresholds.

Conclusions

Both kinds of detection performance analysis showed the method based on Holt-Winters exponential smoothing superior on non-sparse time series with day-of-week effects. The adaptive CUSUM and Shewhart methods proved optimal on sparse data and data without weekly patterns.

Data median	Detection Delay, days	Holt-Winters	Regression EWMA	Adaptive Shewhart	Adaptive CUSUM	7-day Trend-adj. EWMA	7-day Temporal Scan
Median = 0	Median	3	2	4	2	4.5	2
	Penalized Mean	7.2	7	6.6	6.2	7.3	7.6
Median > 0	Median	2	2	2.5	2	6	4
	Penalized Mean	6.1	7	7.2	7.1	7.7	6.8

Keywords

biosurveillance; timeliness; detection; alerting methods; sensitivity

References

1. Elbert Y, Burkom H, Shmueli G, Development and evaluation of a data-adaptive alerting algorithm for univariate temporal biosurveillance data Stat. Med. 2009; 28:3226-3248
2. Burkom H, Elbert Y, Thompson M, et al, Development, adaptation, and assessment of alerting algorithms for biosurveillance JHUAPL Technical Digest, Volume 24, Number 4 (2003)

*Yevgeniy Elbert

E-mail: yevgeniy.elbert@jhuapl.edu

