# Resolving the 'Boy Who Cried Wolf' Syndrome

**Michael A. Coletta, Julie Plagenhoef, Howard S. Burkom, Wayne Loschen**
*Virginia Department of Health, Johns Hopkins Applied Physics Lab*

### OBJECTIVE

**This study's aim was to refine an automated biosurveillance system in order to better suit the daily monitoring capabilities and resources of a health department.**

### BACKGROUND

To date, most syndromic surveillance systems rely heavily on complicated statistical algorithms to identify aberrations. The assumption is that when the statistics identify something unusual, follow-up should occur. However, with multiple strata analyzed, small numbers for some strata, and wide variances in daily counts, the statistical algorithms will generate flags too often. Experience has shown that these flags usually have little or no public health significance. As a result, syndromic surveillance systems suffer from the 'boy who cried wolf' syndrome. It is clear that the analyst's ability to use professional judgment to sift through multitudes of flags is very important to the success of the system, which suggests that statistics alone cannot identify issues of public health importance from ED data.

### METHODS

This study looks at changing the paradigm in which the statistics operate by providing the user with selective control of the thresholds according to local requirements, experience, and judgment. Two independent epidemiologists reviewed time series plots across 6 syndromes, 4 geographic areas, and 5 age groups. For each stratum, minimum counts for alerting were identified using only professional judgment. Each epidemiologist looked at the variability in the time series graphs with 2 years worth of data and chose a minimum count of interest for each syndrome/region/age combination. These minimum values were cataloged in a table for incorporation in the statistical algorithms. Alerting thresholds for the test statistics were also selectively modified. Statistical flags were suppressed unless both the modified alert threshold was tripped and the minimum count was exceeded. In addition, the study investigated the utility of separate automated warnings based on the appearance of key individual chief complaints chosen by the local epidemiologist. Key words, phrases, and combinations of variables were used to train the natural language processor (NLP) on which records have potential interest to epidemiologists/operators. For example, a record listing a chief complaint of bloody diarrhea or a record representing a young adult admitted to the hospital might be of public health interest. These individual records of potential interest were then presented to the epidemiologist/operator in a separate queue for review.

### RESULTS

A comparison of traditional statistical flags versus user-controlled results tested whether (1) controlling the minimum number of counts and alerting thresholds needed to cause a flag and (2) training the NLP to return a list of key chief complaints can successfully enhance the syndromic surveillance system's ability to alert users of potential events of public health concern.

### CONCLUSIONS

Automated alerting methods must be appropriate for the level and scope of the surveillance activities of the users, and for the geographic range and demographic characteristics of the monitored population. For effective surveillance operations, automated systems should provide knowledgeable, experienced users with the capability to selectively adjust sensitivities according to their needs and resources.

Further Information:
Michael Coletta, michael.coletta@vdh.virginia.gov
Julie Plagenhoef, Julie.Plagenhoef@vdh.virginia.gov
Phone: 804-864-8141