Syndromic Surveillance: STL for Modeling, Visualizing, and Monitoring Disease Counts

David Anderson, PhD

Xavier University of Louisiana, Department of Mathematics, New Orleans-LA

ABSTRACT

Background: Public health surveillance is the monitoring of data to detect and quantify unusual health events. Monitoring pre-diagnostic data, such as emergency department (ED) patient chief complaints, enables rapid detection of disease outbreaks. There are many sources of variation in such data; statistical methods need to accurately model them as a basis for timely and accurate disease outbreak methods.

Methods: Our new methods for modeling daily chief complaint counts are based on a seasonal-trend decomposition procedure based on loess (STL) and were developed using data from the 76 EDs of the Indiana surveillance program from 2004 to 2008. Square root counts are decomposed into inter-annual, yearly-seasonal, day-of-the-week, and random-error components. Using this decomposition method, we develop a new synoptic-scale (days to weeks) outbreak detection method and carry out a simulation study to compare detection performance to four well-known methods for nine outbreak scenarios.

Result: The components of the STL decomposition reveal insights into the variability of the Indiana ED data. Day-of-the-week components tend to peak Sunday or Monday, fall steadily to a minimum Thursday or Friday, and then rise to the peak. Yearly-seasonal components show seasonal influenza, some with bimodal peaks. Some inter-annual components increase slightly due to increasing patient populations. A new outbreak detection method based on the decomposition modeling performs well with 90 days or more of data. Control limits were set empirically so that all methods had a specificity of 97%. STL had the largest sensitivity in all nine outbreak scenarios. The STL method also exhibited a well behaved false positive rate when run on the data with no outbreaks injected.

Conclusion: The STL decomposition method for chief complaint counts leads to a rapid and accurate detection method for disease outbreaks, and requires only 90 days of historical data to be put into operation. The visualization tools that accompany the decomposition and outbreak methods provide much insight into patterns in the data, which is useful for surveillance operations.