Initial evaluation of the early aberration reporting system Florida

Morbidity and Mortality Weekly Report, August 26, 2005 by Yiliang Zhu, W. Wang, D. Atrubin, Y. Wu

Abstract

Introduction: In recent years, many syndromic surveillance systems have been deployed around the United States for the early detection of biologic terrorism-related and naturally occurring outbreaks. These systems and the associated aberration detection methods need to be evaluated.

Objective: This study evaluated several detection methods of the Early Aberration Reporting System (EARS) under serially correlated syndromic data and to demonstrated the need for calibrating these methods.

Methods: In an initial evaluation of the Syndromic Tracking and Reporting System in Hillsborough County, Florida, serially correlated syndromic data were simulated using statistical models in conjunction with real syndromic data. The detection methods were tested against two patterns of simulated outbreaks. They were compared using a conditional average run length and a receiver operating characteristic curve under defined patterns of detection.

Results: Increasing serial correlation inflates the false alarm rate and elevates sensitivity. Among the detection methods in EARS, C2 and P-chart have the best overall receiver operating characteristic curve within the context of the simulations. C2 is least affected by the serial correlation, the outbreak type, and the defined patterns of detection signal.

Conclusion: Evaluation of the detection methods needs to be adaptable to the constantly changing nature of syndromic surveillance. Deployment of EARS and other methods requires adjusting the false alarm rate and sensitivity in accordance with the syndromic data, the operating resources, and the objectives of the local system. For timely detection, C2 is

superior to other methods, including C3, under the simulation conditions. P-chart is the most sensitive when the serial correlation is negligible.

Introduction

The risk for biologic terrorism attacks has promoted the development and deployment of syndromic surveillance systems in the United States and around the world (1). The majority of these systems use patient or consumer encounter data from multiple sources (e.g., hospital emergency departments (ED), military facilities, or theme parks). Workplace absenteeism and over-the-counter drug sales are also being monitored for statistical aberrations. The data are converted to specified syndrome categories and analyzed to detect significant temporal or spatial aberrations that deviate from the expected baseline trends. Although some systems conduct spatial analysis (2), most use trend analysis to detect temporal aberrations.

Developed by CDC, the Early Aberration Reporting System (EARS) consists of a class of quality-control (QC) charts, including Shewhart chart (P-chart), moving average (MA), and variations of cumulative sum (CUSUM) (3). Many syndromic surveillance systems use EARS for temporal aberration detection (4); some also use other QC charts such as exponentially weighted moving average (EWMA) (5,6). A common characteristic in adopting these Qc charts for syndromic data analysis is the use of a sample estimate for the baseline mean and standard deviation (SD). This approach circumvents the difficulties associated with the modeling of the baseline trend of the syndrome, a process complicated by the discreteness, serial correlation, seasonality, and daily fluctuation of the syndromic data. At present, understanding of these methods within the context of syndromic data is limited, and systematic evaluations of syndromic surveillance have not been conducted.

In Qc settings, serial correlations can substantially affect the time length to the first aberration signal (Average Running Length [ARL]) when using such methods as CUSUM and P-chart (7,8). How serial correlation also could affect other performance measures such as the false alarm rate (one minus specificity) or the sensitivity (defined generally as the probability of successful detection [PSD] associated with a pattern of signals) is unclear. Certain signal patterns or event detection (e.g., signaling three days in a row) might provide more information than a single day signal about the strength or duration of the outbreak, thus guiding public health agencies in designating the necessary follow-up investigation according to the strength of the signals (9). Simulation studies were conducted to evaluate the sensitivity and specificity of a single day signal for three EARS' variations of CUSUM (C1, C2, and C3) (/0), a seasonally adjusted CUSUM, and an historic limits method (11). However, these simulations assumed serially independent data, with the magnitude of aberrations ranging between 1.65 to 51.0 times the baseline mean (10) or 2-3 times the baseline standard deviation. Seasonality was added to simulated baseline data in the second study (11), but was not modeled in the baseline mean. The simulation indicated that C3 is superior to C2, without controlling for specificity.

This study focuses on the impact of serial correlation on the PSD of selected patterns of signals and compares the performance of five common detection algorithms (P-chart, C2, C3, MA [employed in EARS], and EWMA [a general statistical QC chart not included in EARS]). Two different outbreak patterns are considered here. The comparisons are based on two criteria. The first criterion is a conditional average run length (CARL) given successful detection in a given time window. The second is the receiver operation characteristic (ROC) curves. The purpose is to provide guidance for selection of detection methods and to illustrate the need for calibrating the methods to attain a required level of specificity and sensitivity.