Choice of Baseline Period for Limited Baseline Aberration Detection Methods

Ronald E Gangnon, Ph.D., Marc A. Bellazzini, M.D.
School of Medicine and Public Health, University of Wisconsin–Madison

OBJECTIVE

This paper evaluates the operating characteristics of limited baseline aberration detection methods using different lengths (7-28 days) and end dates (1-7 days prior to the current day) for the baseline period using simulated outbreaks added to real data and simulated data representative of real data.

BACKGROUND

Three limited baseline aberration detection methods, C1, C2 and C3 [1], have been widely implemented by public health departments as a tool for on-going syndromic surveillance [2]. All three methods use a moving 7-day window as a baseline period for estimation of the mean and standard deviation (SD) of the current day's case count. In this paper, we evaluate the impact of window length and end date of the baseline period on the operating characteristics of this type of limited baseline aberration detection method.

METHODS

C1 and C2 compare the observed case count on the current day to the mean for the 7-day baseline period, ending 1 or 3 days prior to the current day, respectively. In either case, values that exceed the baseline mean by a fixed number of SDs, typically 3, are flagged as aberrations. Here, we varied the length of the baseline period varies from 7 to 28 days and the last day of the baseline period from 1 to 7 days prior to the current day. Counts were square-root transformed to better approximate normality. Thresholds for alerts, sqrt(1+1/d)*qt(1-FD,d-1)*SD, were chosen to achieve a desired nominal false detection rates (FD) using the *t* distribution rather than the standard normal to account for sampling error in the estimated baseline mean and SD.

We obtained daily counts of two syndromes, influenza (ICD-9 discharge code) and fever (recorded temperature above 100.4 F), in the Emergency Department of the University of Wisconsin Hospital and Clinics from June 13, 2007 to June 11, 2008. To estimate false detection rates, we generated 10,000 simulated daily counts (independent Poisson or negative binomial random variables) representative of the number of cases of influenza (mean 6.02, SD 9.52) and fever (mean 2.34, SD 3.37) observed in our data to represent the in-control (non-aberration) process. To estimate true detection rates, we added an aberration, a Poisson random variable with mean 5, to each of the last 333 days of the observed processes or 10,000 simulations from the in-control process.

RESULTS

All methods were well-calibrated, e.g. false detection rates were at or below the nominal levels. For both real and simulated datasets, true detection rates for methods were substantially higher for methods based on longer windows.

	7-day		14-day		28-day	
Scenario	FD	TD	FD	TD	FD	TD
Flag if daily count exceeds threshold based on nominal 5% FD rate						
Poi(2.3)	3.4%	54.1%	3.1%	62.6%	2.7%	66.5%
NB(2.3,3.4)	3.3%	44.3%	3.1%	51.4%	3.2%	55.3%
Fever[a]	*3.6%*	49.3%	*3.9%*	56.1%	*2.7%*	59.1%
Poi(6.0)	4.2%	41.3%	4.2%	47.6%	4.2%	50.2%
Poi(6.0,9.5)	4.7%	30.0%	4.6%	34.1%	4.4%	36.6%
Influenza[a]	*4.5%*	37.1%	*2.4%*	43.0%	*2.7%*	46.3%

Table: False detection (FD) and true detection (TD) rates for methods with baseline periods of 7, 14 and 28 days starting the day prior to the current day based on actual and simulated data representative of daily counts of fever and influenza in UW Hospital and Clinics Emergency Department, June 2007-08. [a]FD rate is the proportion of flagged days in last 333 days of actual data and reflects an unknown proportion of true and false detections.

CONCLUSIONS

Our results suggest that (1) the use of thresholds based on the *t* distribution for aberration detection controls the false detection rate and (2) the use of longer baseline periods for the estimation of means and SDs is associated with substantial improvements in performance of aberration detection methods. The latter finding is in agreement with earlier studies with uncalibrated or empirically calibrated procedures [3,4].

REFERENCES

[1] Centers for Disease Control and Prevention. Sundromic surveillance for bioterrorism following the attacks on the World Trade Center–New York City, 2001. MMWR 2002 51: 13-15.

[2] Hutwagner LC, Thompson WW, Seeman GM, Treadwell T. A simulation model for assessing aberration methods used in public health systems used in public health surveillance for systems with limited baselines. StatinMed, 2005: 24: 543-550.

[3] Xing J, Burkom H, Copeland J, Bloom S, Hutwagner L, Tokars J. Performance Characteristics of Control Chart Detection Methods. Adv Dis Surveillance 2007; 4:122.

[4] Chang H, Chen J, Cochrane DG, Allegra JR, Burkom H, Tokars JI, Smith PF. A Pilot Study of Aberration Detection Algorithms with Simulated Data. Adv Dis Surveillance 2007; 4:240.

Further Information:
Ronald Gangnon, ronald@biostat.wisc.edu