

Multiregional CUSUM Surveillance

Peter A. Rogerson, Minakshi Trivedi, Sharmistha Bagchi-Sen

Departments of Geography and Biostatistics, Dept. of Marketing, and Dept. of Geography, University at Buffalo

OBJECTIVE

This paper suggests and illustrates several approaches to surveillance when data are available for several regions.

BACKGROUND

Monitoring sales of over-the-counter products is becoming increasingly common for purposes of public health surveillance [1, 2]. Sales data for anti-diarrheal medications have been used to monitor outbreaks of waterborne *Cryptosporidium* outbreaks [3, 4]. An attractive feature of [5] is its focus upon coupling predictions of sales for a given day (based upon times series methods) with control chart methods from the field of statistical process control [6].

METHODS

Daily sales of *Immodium*, *Pepto Bismol*, and *Kaopectate* were available for several large suburban supermarkets for 2005, based upon data linked with shopper's club cardholders.

Daily counts were initially assumed to come from a Poisson distribution. Data were transformed to normality and adjusted for any overdispersion. Cumulative sum methods [6] were used to monitor deviations between daily counts and expectations. Several multivariate cusum approaches were adopted to account for the simultaneous monitoring from many stores.

RESULTS

Preliminary results for one store revealed that the mean number of daily sales to households was 1.9, and the variance was 2.43.

The cusum threshold was exceeded twice. The most pronounced signal occurred on May 15th and continued above threshold through May 22nd. Counts were relatively higher than average beginning on May 11th, and these counts led up to the signal on May 15th. The mean of the 12 counts for May 11 – May 22 was 3.67, in contrast with the mean of 1.9 observed for the entire dataset.

CONCLUSIONS

Next steps include developing a model of expectations that is based upon day-of-week, monthly, seasonal and holiday effects that may exist. In addition, daily sales may depend upon whether there are promotions associated with products. The dataset contains information that will allow the modeling of these effects. Additional effort will also be given to developing alternative approaches to multiregional surveillance.

REFERENCES

- [1] Magruder, S, Happel Lewis, S, Najmi, A, and Florio, E. 2004. Progress in understanding and using over-the-counter pharmaceuticals for syndromic surveillance. *MMWR* 53, 117-122.
- [2] Wagner, M, Espino, J, Hersh, J et al. 2004. National Retail Data Monitor for public health surveillance *MMWR* 53 (supplement) 40-42.
- [3] Proctor, ME, Blair, KA, and David, JP. 1998. Surveillance data for waterborne illness detection: an assessment following a massive waterborne outbreak of *Cryptosporidium* infection. *Epidemiolog. Infection.* 120(1) 43-54.
- [4] Stirling, R, Aramini, J, Ellis, A et al. 2001. Waterborne cryptosporidiosis outbreak, North battleford, Saskatchewan, Spring 2001. *Can Commun Dis Rep.* 27 (22) 185-192.
- [5] Goldenberg, A, Shmueli, G, Caruan, RA, and Fienberg, SE. 2002. Early statistical detection of anthrax outbreaks by tracking over-the-counter medication sales. *Proc. of the Nat. Acad. of Sciences* 99, 8: 5237-5240.
- [6] Montgomery, DC. 1996. *Introduction to statistical process control*. New York: Wiley.

Further information: Peter A. Rogerson,
rogerson@buffalo.edu