# An Empirical Study of the Effect of Sentinel Sample Size in Syndromic Surveillance Using a Space-Time Permutation Method

**Richard Brand[1], Jie Huang[2], Ken Kleinman[3], Richard Platt[3], John Hsu[2].**

*[1]Department of Epidemiology and Biostatistics, University of California, San Francisco, [2]Division of Research, Kaiser Permanente, Northern California, [3]Department of Ambulatory Care and Prevention, Harvard Medical School and Harvard Pilgrim Health Care*

### *OBJECTIVE*

Our goal was to assess the impact of sentinel sample size and criteria for a signal on performance of daily prospective space-time permutation detection by comparing results in varying size random samples from a large health plan to results found in the full membership.

### BACKGROUND

There is limited closed-form statistical theory to indicate how well the prospective space-time permutation scan statistic will perform in the detection of localized excess illness activity. Instead, detection methods can be applied to simulated data to gain insight about detection performance [1]. Such results are dependent on the way outbreaks are simulated and the nature of the background data. As an alternative, we explore an empirical approach in which the membership of a large health plan is used to represent a community and detection performance is assessed in samples from the larger group.

### METHODS

We applied the prospective space-time permutation method implemented in SaTScan [2,3] to 612,567 respiratory illness episodes among 3.6M Kaiser Permanente members in 405 zip codes in Northern CA in 2005. Space-time clusters detected with a Recurrence Interval (RI) of 1000 or more days (i.e. p<=0.001), were defined as signals in the full membership. These were then organized into sequences of signals that overlapped in time and space. The signal sequences in the full membership were used as a reference to which results from samples were compared. Ten random samples of each size 50K, 100K, 250K and 1M were drawn from the 3.6M members. Signals in the samples were defined using various RI criteria; 20, 30, 90, 180, 365, and 1000 days. As a function of sample size and sample signal criteria, we assessed the detection performance of the samples in terms of sensitivity, specificity, positive predictive value, negative predictive value, timeliness, percentage of reference outbreaks detected, and density of sample signals during the reference sequences.

### RESULTS

In the full 3.6M membership, we found 151 signals with RI>=1000 days. These formed into 14 separate sequences of overlapping signals. These reference sequences varied greatly with respect to location in the coverage area of the health plan, numbers of alerts (1 to 62), number of zip codes involved (1 to 191), days from start of the first signal till end of the last (7 to73), and days of detection lag from the start of first signal in the sequence to the first detection date (5-13) with mean of 9.5 days. Mean performance results for 10 samples of size 250K illustrate the findings for samples. In order to achieve a sensitivity of 14% it was necessary to reduce the signal criterion in the sample to RI>=20 days (p=0.05). In this situation, the specificity was still relatively high at 93%. Despite the low sensitivity, 13 of the 14 reference sequences were hit by sample signals at some time during their course. However, the signals in the sample occurred with an additional detection lag of 4 days on average compared to the detection in the full membership. The other consequence of low sensitivity was a reduction of density of signals during the reference sequences from 1 every 2.2 days with the full membership to one every 9.1 days, on average with samples of 250K and RI >=20.

### CONCLUSIONS

Even when the criterion for an signal is weakened, the sensitivity for detection of illness activity is substantially diminished as sample size decreases. In this empirical study, the low sensitivity mainly expressed itself in additional detection lag and a reduction in the density of signals.

### REFERENCES

[1] Mandl KD, Reis B, Cassa C, Measuring Outbreak-Detection Performance By Using Controlled Feature Set Simulations. MMWR Supplement, September 24,2004/Vol. 53, 130-136

[2] Kulldorff M. Prospective time periodic geographic disease surveillance using a scan statistic. Journal of the Royal Statistical Society, Series A 2001;164:61-72.

[3] Kulldorff M. SaTScan: Software for the spatial and space-time scan statistics. In: www.satscan.org; 2004.

.