COACTION PUBLISHING

# Comparing methods for sentinel surveillance site placement

**Geoffrey Fairchild*, Alberto Segre, Gerard Rushton, Eric Foster and Philip Polgreen**

Department of Computer Science, University of Iowa, Iowa City, IA, USA

## Objective

To evaluate the performance of several sentinel surveillance site placement algorithms for influenza-like illness (ILI) surveillance systems. We explore how these different approaches perform by capturing both the overall intensity and timing of influenza activity in the state of Iowa.

## Introduction

ILI data are collected via an Influenza Sentinel Provider Surveillance Network at the state level. Because participation is voluntary, locations of the sentinel providers may not reflect optimal geographic placement. This study analyzes two different geographic placement schemes—a maximal coverage model (MCM) and a K-median model, two location-allocation models commonly used in geographic information systems (GIS) (1). The MCM chooses sites in areas with the densest population. The K-median model chooses sites, which minimize the average distance traveled by individuals to their nearest site. We have previously shown how a placement model can be used to improve population coverage for ILI surveillance in Iowa when considering the sites recruited by the Iowa Department of Public Health (IDPH) (2). We extend this work by evaluating different surveillance placement algorithms with respect to outbreak intensity and timing (i.e., being able to capture the start, peak and end of the influenza season).

## Methods

We developed a web-based site placement calculator to aid public health officials in designing their surveillance system. We then compare the two algorithmic site placement schemes against each other by simulating the spread of influenza across the state of Iowa. Our simulations are based on an Iowa Medicaid dataset comprised two million cases classified with 30 different ILI ICD-9 codes and their corresponding geocodes from 2000 to 2008. We use the Huff Model (3) to determine whether or not a case might have been detected by a particular network of sites. Using this scheme, we compare surveillance networks based on outbreak intensity (i.e., which networks detect the highest percentage of cases) and outbreak timing (i.e., which networks detect cases temporally in sync with the true start, peak and end of the disease season). To compare network outbreak timing, we generate the noise parameters of a state space time series model using the expectationmaximization (EM) algorithm implementation provided by Shumway and Stoffer (4). We then perform an analysis on which ICD-9 codes a network might consider.

## Results

Considering outbreak intensity, we show that sites chosen by our approach outperform the sites used by the IDPH. In other words, we can provide a more accurate representation of disease burden during an influenza season. However, our approach does not provide a substantial difference in the detection of the start, peak and end of the influenza season. In addition, when using the noise values generated by the EM algorithm to analyze the minimal number of sites needed to estimate the timing of the influenza season, our results were highly influenced by both the number of different ICD-9 codes considered and the number of cases considered.

## Conclusions

We built a web-based tool to assist public health officials in designing their sentinel surveillance site network. Through simulation, we show that the sites our tool selects allow for better representation of disease burden. We also show that selecting the correct ICD-9 codes that the surveillance network should consider may be as important as selecting the locations of the sites themselves. Using our methods, we can help public health officials design surveillance systems, which are smaller and more easily managed but still detect cases that reflect reliable estimates of both disease burden and timing.

## Keywords

Influenza; surveillance; simulation; time series; geographic information systems

## References

1. Church R. Location modelling and GIS. In: Longley P, et al., eds. Geographical information systems: principles, techniques, management and applications, vol. 1. New York City: John Wiley & Sons; 1999. p. 293–303.
2. Polgreen PM, Chen Z, Segre AM, Harris ML, Pentella MA, Rushton G. Optimizing influenza sentinel surveillance at the state level. Am J Epidemiol. 2009;170:1300–6.
3. Huff D. A probabilistic analysis of shopping center trade areas. Land Econ. 1963;39:81–90.
4. Shumway RH, David SS. Time series analysis and its applications: with R examples. 2nd ed. New York: Springer Science+ Business Media; 2006.

***Geoffrey Fairchild**
E-mail: geoffrey-fairchild@uiowa.edu



Citation: Emerging Health Threats Journal 2011, **4**: 11145 - DOI: 10.3402/ehtj.v4i0.11145
(page number not for citation purpose)