Detecting Outbreaks and Other Clusters in Reportable Disease Data

Shannon L. Page, BS, Richard M. Thomas, BS,
Brian E. Fowler, MPH, Lynn K. Giljahn, MPH
The Ohio Department of Health

OBJECTIVE

Clusters of cases of individually-reportable infectious diseases were identified by a spatial-temporal retrospective analysis. Clusters were examined to determine association with previously reported outbreaks.

BACKGROUND

Outbreaks of infectious diseases are identified in a variety of ways by clinicians and public health practitioners but not usually by analytic methods typically employed in syndromic surveillance. Systematic spatial-temporal analysis of statewide data may enable earlier detection of outbreaks and identification of multi-jurisdictional outbreaks.

METHODS

Analysis focused on diseases with gastrointestinal symptoms that commonly cause outbreaks: Salmonellosis, Shigellosis, Campylobacteriosis, Cryptosporidiosis, Giardiasis, and *Escherichia coli*. Data from reported outbreaks and data on individual cases of diseases that met CDC case definition and were in Ohio's system for infectious disease surveillance, Ohio Disease Reporting System (ODRS), were used for this analysis. Data were aggregated by earliest date associated with each case and ZIP code of residence. Year 2000 U.S. Census data were used for population denominators. The analysis included individual cases of disease and outbreaks for the period January-June 2008. Poisson time-space modeling was used with a maximum cluster size of 20 km and a maximum time frame between cases of 7 days[1]. Clusters of individual cases of disease with p-values of less than 0.05 were investigated further to determine association with previously reported outbreaks. Individual cases from clusters not associated with known outbreaks were reviewed further for signs of common exposure. Each cluster was further characterized spatially and temporally with an epidemiologic curve and a map[2] of individual cases (Figure 1). Data for this project were aggregated and compiled using SAS 9.1 (SAS Institute Inc., Cary, NC).

RESULTS

Analysis yielded 27 statistically significant clusters. Of these, 13 clusters were associated with 13 known outbreaks. Five known outbreaks were not detected.

One of the 27 clusters not associated with a known outbreak indicated an unusually high number of Campylobacteriosis cases in a two-county, rural area. To date, local public health investigation has not identified a common exposure between these cases.
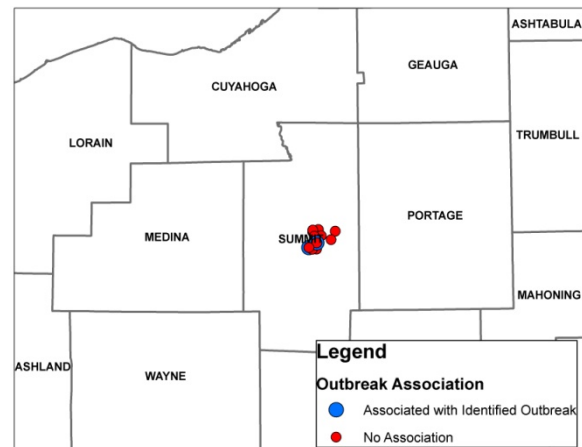

Figure 1: A statistically significant cluster of Shigellosis cases in the greater Akron, Ohio area with outbreak association status.

CONCLUSIONS

Preliminary results suggest that clusters with the highest statistical significance were often associated with previously identified outbreaks. Ongoing, prospective analysis may be useful for detecting unidentified outbreaks and disease clusters across jurisdictional boundaries. Timeliness of such analysis may be problematic, in that the data used rely heavily on the promptness of reporting of disease by laboratories and local partners. In some cases, further investigation of case clusters by public health may be warranted.

Cluster analysis will be expanded to include other reportable conditions as analysis criteria are refined.

REFERENCES

1. SaTScan, version 7.0.3, Information Management Services, Inc.
2. ArcMap, version 9.1, ESRI