CLUSTER DETECTION INCORPORATING LAGGED TEST DATA

Kevin Konty MS and Farzad Mostashari MD M.Sc

Bureau of Epidemiology Services, New York City Department of Health and Mental Hygiene

OBJECTIVE

Using New York City's dead bird surveillance for West Nile Virus (WNV), this paper presents two explorations of the spatial cluster detection problem in which lagged test results are available for a random subset of observations. First, we establish a framework for the direct evaluation of methods and identify the optimal parameterization over a large family of models. We then investigate ways in which the lagged test results and other covariates might be used prospectively to extend the family of models by refining the baseline.

BACKGROUND

Spatial cluster detection methods using empirical syndromic data are extremely difficult to evaluate. An ideal framework for evaluation would include testing of the points within the system. Advances in public health informatics may eventually provide spatially-referenced lab data; some linked directly to the syndromic observations. WNV dead bird surveillance already has this capacity.

A number of techniques have been proposed to identify areas of elevated WNV risk including dead crow densities and simple tests for spatial pattern [2-3]. While these methods attempt to identify areas where humans are at elevated risk, the null hypothesis is that there is no WNV in the bird population and so clusters identified presumably occur because some of its member birds are infected. The link to humans is secondary and while it must also be established, the system cannot work if it does not identify areas of elevated bird risk. In New York City, dead bird patterns reported are prospectively searched for clusters. One question arising is how to identify the spatial pattern under the null hypothesis. Using recent data may result in WNV+ birds in the baseline, while using early-season patterns ignores the spatial dynamics of bird mortality.

METHODS

We employ the general framework for dead bird surveillance described by Mostashari[1], used since 2001 in New York City. Spatial clusters are identified using spatial case/control SaTScan. The clusters identified represent areas of elevated WNV Risk. The parameters in this setting include threshold p-values, bird species used, and time period specifications for the cases (n days) and baseline (t_1,t_2). A random subset of birds is tested and used for evaluation. The logic is simple: tested birds from clusters should have a higher prevalence of WNV than those outside of clusters. Because of the nature of SaTScan clusters, we might also expect that positive birds outside of clusters tend to be closer in space and time to clusters than negative birds; or that they appear in less significant clusters.

Theoretically, the null hypothesis would best be constructed from WNV-negative birds. To approximate this, test results were used to estimate the probability of infection for untested birds. We explore several methods including smoothing methods and simple logistic modeling incorporating species, seasonality, and mosquito data. The resulting baseline pattern is a series of points with probabilities that can be analyzed using continuous distribution SaTScan (from SaTScan 7).

RESULTS

Previously, New York City used early season patterns defined as those bids reported between April 1 and the date of first positive test result. In 2003 and 2004 this method resulted in clusters with mildly elevated prevalence of WNV, while in 2005 the clusters did not predict positive birds at all. Using a baseline of April 1 to 2 weeks before the date of analysis slightly improved results for all years (2005 rr=1.68). However, using a baseline for the trailing 3 weeks drastically improved the clusters ability to identify positive birds, with a 2005 relative risk of 3.9. Reduction of bids species provided little benefit.

Modeling of WNV infection in birds using covariates and test results will take place over the 2006 WNV season. An evaluation of the models as well as risk maps will be presented.

CONCLUSIONS

The direct evaluation of spatial cluster methods can greatly improve surveillance. For West Nile Virus, this approach has substantially improved identification of areas with WNV risk. The methods presented here are applicable to situations with random testing of a subset of observations; the direction in which syndromic surveillance is headed.

REFERENCES

[1] Mostashari F et al. "Dead bird clusters as an early warning system for West Nile virus activity." EID 2003 Jun.

[2] Eidson M et al. "Dead crow density and West Nile virus monitoring, New York." EID 2005 Sep

[3] Theophilides CN, et al "West Nile virus risk areas: the Dynamic Continuous-Area Space-Time system." AJE 2003 June

Further Information: Kevin Konty, kkonty@health.nyc.gov