Using Space-Time Scan Statistics to Detect Pertussis and Shigellosis Outbreaks

Jian-Hua Chen MD, MsPH, Charlene Weng MS, Hwa-Gan Chang PhD

New York State Department of Health, Albany, New York

CORRESPONDING AUTHOR

Jian-Hua Chen, MD, MsPH
Statistical Unit, Division of Epidemiology
New York State Department of Health
Corning Tower Rm 736, ESP
Albany, NY 12237
Phone: 518-402-5012 FAX: 518-474-4880
E-mail: jhc04@health.state.ny.us

BACKGROUND

➤ Early detection of disease outbreaks is important to minimize morbidity and mortality through timely implementation of disease prevention and control measures.

➤ Scan statistics can detect disease clusters that are associated in space, time, or space-time, indicating potential disease outbreaks.

➤ New York State cases of communicable diseases are required to be reported to health department with information of disease date and patient's geographic location, which can be analyzed with scan statistics.

OBJECTIVES

➤ Apply scan statistics to pertussis and shigellosis cases reported to the New York State Department Communicable Disease Electronic Surveillance System (CDESS) to evaluate its utility in detection of disease outbreaks.

➤ Develop a method to easily run the scan statistics analysis and reporting, which will allow conduct routine data analysis and early investigation of potential outbreaks.

METHODS

➤ The prospective space-time permutation model of SatScan (a freely available scan statistics software) was used to identify space-time cluster of pertussis and shigellosis reported to CDESS in 2012 (excluding NY City cases).

➤ The case investigation date was used as the time variable, and patient's zip as the geographic location variable in the scan statistics analysis. Maximum detectable cluster size was set as a circle in <= 20 kilometer (km) radius that could span up to 15 days.

➤ Twenty-four (24) separate space-time scan statistics analyses were performed (two for each month of 2012) with 2011 cases as reference.

➤ A SAS program was developed to carry out the scan statistics analyses, which also produced a map of detected clusters in zip code, cluster parameters table, and case listing of the cluster.

➤ The outbreak investigation status of the reported cases in CDESS was used to evaluate the accuracy of clusters detected by the scan statistics. For the purpose of this study, a cluster was considered as confirmed outbreak if it contains one or more outbreak case(s).
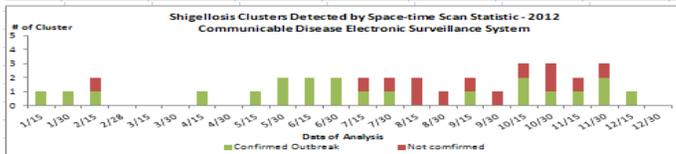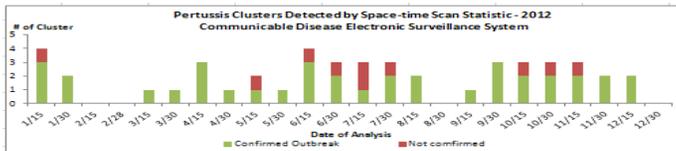
RESULTS

There were 2,719 pertussis and 802 shigellosis cases in NYS (excluding NYC) reported in 2012. At the p < 0.05 significance level, the scan statistics identified 47 pertussis and 34 shigellosis clusters, of which 37 and 21 were confirmed as outbreaks, yielded positive predictive values (PPV) of 79% and 62% respectively.

Clusters detected by scan statistic (space-time premutation model)
Communicable Disease Electronic Surveillance System

Disease	Number of clusters detected	Clusters confirmed as outbreaks	Positive Preditive Value	Median Case Counts in Clusters (range)	Median Areas (km) of Clusters (range)	Median Duration (days) of Clusters (range)
Pertussis	47	37	78.7%	6 (2-19)	11.6 (0-20)	7 (1-15)
Shigellosis	34	21	61.8%	8 (3-29)	9.2 (0.1-18.5)	10 (1-14)

Pertussis Clusters Detected by Space-time Scan Statistic - 2012
Communicable Disease Electronic Surveillance System

Shigellosis Clusters Detected by Space-time Scan Statistic - 2012
Communicable Disease Electronic Surveillance System

Map of Pertussis Clusters by Patient's ZIP Code, ending 2012-03-15

Table of Pertussis Clusters by Patient's ZIP Code
Study Start = 2011/01/01
Study End = 2012/03/15
Maximum Cluster Radius in km = 20 km
Maximum Cluster in Days = 15 day(s)

Cluster	P_Value	Reccurance Interval (days)	Latitude	Longitude	Radius (km)	Start_Date	End_Date	ZIP_Code	Observed_#	Expected_#
1	0.0013	747	42.5058	-76.6897	0	2012/3/8	2012/3/15	14886	5	0

Total ZIP Codes = 1
Total Observed Patients = 5
Total Expected Patients = 0

Obs	disease	CLUSTER	InvestCreate_date	county_case	zipcode	status_case	outbreak
1	Pertussis	1	03/08/2012	TOMPKINS	14886	probable	outbreak
2	Pertussis	1	03/12/2012	TOMPKINS	14886	probable	outbreak
3	Pertussis	1	03/12/2012	TOMPKINS	14886	probable	outbreak
4	Pertussis	1	03/12/2012	TOMPKINS	14886	probable	outbreak
5	Pertussis	1	03/15/2012	TOMPKINS	14886	probable	outbreak

CONCLUSIONS

➤ The space-time permutation scan statistics only requires event counts, date, and location, which can be utilized to detect disease clusters with disease surveillance data routinely collected by public health agency.

➤ The scan statistics analysis can be set up as a routine procedure to identify possible outbreaks early with information of cluster size, time, and location to facilitate further outbreak investigation.

➤ The current study showed that scan statistics is a useful tool for detecting pertussis and shigellosis outbreaks with reasonable PPVs.

➤ It is important to set up the scan parameters properly according the disease of interests to achieve optimal results as the scan statistics result can be affected by disease patterns such as case frequency, incubation period, transmission pattern, etc..