Major Article

Implementation and evaluation of an automated surveillance system to detect hospital outbreak

Anna Stachel MPH, CIC [a,*], Gabriela Pinto BS [a], John Stelling MD, MPH [b], Yi Fulmer PhD [c], Bo Shopsin MD, PhD [c], Kenneth Inglima MS [d], Michael Phillips MD [a,c]

[a] Infection Prevention and Control, New York University Langone Health System, New York, NY
[b] School of Medicine, Brigham and Women's Hospital, WHO Collaborating Centre for Surveillance of Antimicrobial Resistance, Boston, MA
[c] Division of Infectious Diseases & Immunology, New York University Langone School of Medicine, New York, NY
[d] Clinical Microbiology and Diagnostic Immunology, New York University Langone School of Medicine, New York, NY

Background: The timely identification of a cluster is a critical requirement for infection prevention and control (IPC) departments because these events may represent transmission of pathogens within the health care setting. Given the issues with manual review of hospital infections, a surveillance system to detect clusters in health care settings must use automated data capture, validated statistical methods, and include all significant pathogens, antimicrobial susceptibility patterns, patient care locations, and health care teams.
Methods: We describe the use of SaTScan statistical software to identify clusters, WHONET software to manage microbiology laboratory data, and electronic health record data to create a comprehensive outbreak detection system in our hospital. We also evaluated the system using the Centers for Disease Control and Prevention's guidelines.
Results: During an 8-month surveillance time period, 168 clusters were detected, 45 of which met criteria for investigation, and 6 were considered transmission events. The system was felt to be flexible, timely, accepted by the department and hospital, useful, and sensitive, but it required significant resources and has a low positive predictive value.
Conclusions: WHONET-SaTScan is a useful addition to a robust IPC program. Although the resources required were significant, this prospective, real-time cluster detection surveillance system represents an improvement over historical methods. We detected several episodes of transmission which would have eluded us previously, and allowed us to focus infection prevention efforts and improve patient safety.

BACKGROUND

The timely identification of a cluster—a larger than expected number of cases within an established time frame—is a critical requirement for infection prevention and control (IPC) departments because these events may represent transmission of pathogens within the health care setting. A recent survey revealed hospitals typically rely on a periodic visual inspection of line-lists to detect clusters instead of statistically validated methods.[1] Because the ongoing creation of line-lists is laborious, hospitals focus on only a small subset of epidemiologically important organisms, such as methicillin-resistant *Staphylococcus aureus*, carbapenem-resistant *Enterobacteriaceae*, and *Clostridium difficile*.[1] The creation of the line-list is not an end in itself—for this data to be useful, it must be analyzed for statistically significant changes, and potential increases must be systematically evaluated to prevent further infections.[2] The subsequent investigation of a cluster by IPC personnel is complicated because the dynamic health care setting affords multiple routes for the transmission of pathogens: health care worker to patient, environment to patient, and patient or visitor to patient. A detailed investigation is required, given patients colonized or infected with potentially infectious pathogens share the multiple health care personnel, patient care locations, and mobile equipment with many susceptible individuals, and undergo a variety of medical treatments and procedures which increase both vulnerability to subsequent infection and potential for disseminating these pathogens to others.[3] Control of hospital outbreaks requires rapid identification and investigation of clusters, with concurrent implementation of appropriate control measures.

* Address correspondence to Anna Stachel, MPH, CIC, New York University Langone Health System, Greenberg Hall, SC1-132, 545 First Ave, New York, NY 10016.
E-mail address: anna.stachel@nyumc.org (A. Stachel).
Conflicts of interest: None to report.

Given these challenges, we think a surveillance system to detect clusters in the health care settings must use automated data capture, validated statistical methods, and include all significant pathogens, antimicrobial susceptibility patterns, patient care locations, and health care teams. This system should identify clusters in near to real time while accounting for historical background rates of occurrence. In addition to early detection of clusters, comprehensive and validated outbreak surveillance systems may improve our understanding of transmission dynamics and enhance our capability to prevent outbreaks.[1,4,5] The characteristics of such a surveillance system should include the following:

1. An integrated laboratory information system which consistently and routinely provides test characteristics, results, and antimicrobial susceptibilities of microbes in a structured data format.
2. Analytics which link microbiology results with data elements in the electronic health record (EHR) to detect clusters. Additional data elements associated with the patients in a cluster are automated to minimize manual medical record review by IPC personnel.
3. The application of validated statistical methods to quantify the probability of an observed group of organisms is significantly different than expected.
4. Stratification of clusters into levels based on the likelihood a cluster represents transmission of pathogens in the health care setting. This stratification is derived from statistical measures, pathogen type, and patient characteristics, and will allow optimization of IPC resources.
5. The data associated with a cluster are displayed in a comprehensive and easy to use format to facilitate an efficient review by IPC practitioners.
6. Clusters are systematically investigated to determine whether health care transmission of the pathogen was possible, and if so, identify potential routes of transmission and implement control measures. Additional tests may be obtained, such as environmental cultures, bacterial strain typing, and molecular diagnostics.
7. If IPC determines the cluster may indeed represent transmission, the investigation results are reviewed with the clinical team, and an apparent cause analysis is performed to identify causal factors resulting in transmission of pathogens.
8. Summary data are produced and reported to identify trends and opportunities for improvement.

The SaTScan statistical software (freeware developed by the National Cancer Institute, Boston, MA) uses a space-time permutation scan statistic to identify potential clusters, and WHONET software (Brigham and Women's Hospital, Boston, MA) allows the upload and analysis of isolate-level microbiology laboratory data.[6,7] SaTScan coupled with WHONET can be used to create a comprehensive outbreak detection system in hospitals.[8,9]

SaTScan's space-time scan statistic was originally developed for early detection of disease outbreaks, both geographically and temporally.[10,11] The spatial scan statistic imposes a circular window on top of a map and lets the circle change sizes and move over different areas. The temporal aspect of the scan statistic is represented by the height of the circle or cylinder. For each cylinder, the maximum likelihood ratio is calculated using the number observed and expected cases within and outside the cylinder.[10] Advantages to the space-time statistic as a surveillance tool include the following: adjusting for seasonal variations, searching for clusters without specifying their size and avoiding preselection bias, and creating a single P value.[10,12] The space-time scan statistic used with WHONET-SaTScan does not incorporate denominator data (eg, patient days) to fit the typical Poisson or Bernoulli distributions. This modification

was validated in one study where results using numerator data (eg, infections) were consistent with the analyses that took the patient days and temporal trends into account.[8] Therefore, WHONET software has been shown to be a useful tool for the prospective surveillance of hospital data given it is more adaptable and easier to use than SaTScan alone and approximates the analysis well despite the fact it does not take populations-at-risk into account.[8] In addition, WHONET-SaTScan may be further modified to use unit, groups of units, and service line as a proxy for a geographic area. Our goal was to develop a comprehensive outbreak surveillance system using the WHONET-SaTScan tool and evaluate the system using Centers for Disease Control and Prevention's guidelines.[13]

METHODS

Study population

NYU Langone Medical Center is an academic medical system comprised of 2 hospitals with a total of 925 beds at the time our cluster surveillance system was implemented and evaluated. Our IPC microbiology data repository includes patient demographic information, unit, and clinical service at the time of specimen collection and the unit 3 days prior, collection date and specimen source, and hospital admission date. Antimicrobial susceptibility interpretations are based on Clinical and Laboratory Standards Institute standards.[14]

Development of automated outbreak detection surveillance system

In 2014, our IPC department implemented the WHONET-SaTScan software with an Access database (Microsoft, Redmond, WA) to detect, store, and link EHR data and record results from investigations and interventions of clusters to create an automated outbreak detection surveillance system. Using the space-time permutation (simulated prospective) analysis, we analyzed the first positive culture for each patient every 30 days and defined a cluster using the following predefined parameters: a maximal outbreak window of 60 days and baseline data window of 365 days.[4] SaTScan also calculates the recurrence interval, defined as the statistical likelihood that an observed signal could be caused by chance alone. For example, the longer the recurrence interval, the lower the likelihood of the cluster occurring randomly. In this study, we defined a signal worthy of further analysis and investigation as one with a recurrence interval of >20 days, which equates to $P < .05$. Because there is currently no gold standard to identify epidemiologically important clusters using this methodology, we conducted multiple analyses using different populations, patient location, and cultures. The populations analyzed included the entire hospital population, an individual hospital unit, and a hospital service. The patient location was analyzed by location at time of culture and location 3 days prior to culture. Analysis was also performed using all cultures and only cultures collected 3 days after arrival to the hospital or hospital unit. The WHONET-SaTScan default antibiotic susceptibility profiles for the following 11 pathogen groupings that are of the highest epidemiologic importance to our health system included the following: *Staphylococcus* spp, *Streptococcus* spp, *Streptococcus pneumonia*, viridians group streptococci, *Enterococcus* spp, *Salmonella* spp, *Shigella* spp, *Pseudomonas* spp, *Haemophilus* spp, *Campylobacter* spp, and *Neisseria meningitides*. Figure 1 outlines the flow of microbiology data into the IPC database. Electronic microbiology data are converted to a file recognized by WHONET daily using automated BacLink software (Brigham and Women's Hospital, Boston, MA). Using batch mode, WHONET-SaTScan runs on a daily schedule to import and analyze the BacLink file. Cluster results from SaTScan analysis are exported into a folder where it is subsequently imported into a hospital IPC database using a SAS program script (SAS Institute, Cary, NC) written