HHS Public Access

Author manuscript

Infect Control Hosp Epidemiol. Author manuscript; available in PMC 2022 September 07.

Published in final edited form as:

Infect Control Hosp Epidemiol. 2020 September ; 41(9): 1016–1021. doi:10.1017/ice.2020.233.

Automated Outbreak Detection of Hospital-Associated Pathogens: Value to Infection Prevention Programs

Meghan A. Baker, MD, ScD[1,2], **Deborah S. Yokoe, MD, MPH**[3], **John Stelling, MD, MPH**[2], **Ken Kleinman, ScD**[4], **Rebecca E. Kaganov, BA**[1], **Alyssa R. Letourneau, MD, MPH**[2], **Neha Varma, MPH**[1], **Thomas O'Brien, MD**[2], **Martin Kulldorff, PhD**[2], **Damilola Babalola, MD, MPH**[2], **Craig Barrett, Pharm.D., BCPS**[5], **Marci Drees, MD, MS**[6], **Micaela H. Coady, MS**[1], **Amanda Isaacs, MSPH**[1], **Richard Platt, MD, MSc**[1], **Susan S. Huang, MD, MPH**[7],

For the CDC Prevention Epicenters Program

[1]Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA USA

[2]Brigham and Women's Hospital, Boston, MA USA

[3]University of California San Francisco, San Francisco, CA USA

[4]University of Massachusetts Amherst, Amherst, MA USA

[5]Premier, Inc. Charlotte, NC USA

[6]ChristianaCare, Newark, DE USA

[7]University of California Irvine School of Medicine, Orange, CA USA

Abstract

Objective: To assess the utility of an automated, statistically-based outbreak detection system to identify clusters of hospital-acquired microorganisms

Design: Multicenter retrospective cohort study

Setting: 43 hospitals using a common infection prevention surveillance system

Methods: A space-time permutation scan statistic was applied to hospitals' microbiology and admission, discharge and transfer data to identify clustering of microorganisms within hospital locations and services. Infection preventionists were asked to rate the importance of each cluster. A convenience sample of 10 hospitals also provided information about clusters previously identified through their usual surveillance methods.

Results: We identified 230 clusters in 43 hospitals involving Gram-positive and negative bacteria and fungi. Half of the clusters progressed after initial detection, suggesting that early detection could trigger interventions to curtail further spread. Infection preventionists reported that they would have wanted to be alerted about 81% of these clusters. Factors associated with clusters judged to be moderately or highly concerning included high statistical significance, large size, and

Corresponding Author: Meghan Baker, MD, ScD Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, 401 Park Drive, Suite 401 East, Boston, MA 02215, meghan_baker@harvardpilgrim.org, Phone: 617 867 4946.

Conflict of Interest

All authors report no conflicts of interest relevant to this article.

clusters involving *Clostridioides difficile* or multidrug-resistant organisms. Based on comparison data provided by the convenience sample of hospitals, only 18% (9/51) of the clusters detected by usual surveillance met statistical significance and of their 70 clusters not previously detected, 58 (83%) involved organisms not routinely targeted by the hospitals' surveillance programs. All infection prevention programs felt that an automated outbreak detection tool would improve their ability to detect outbreaks and streamline their work.

Conclusions: Automated, statistically-based outbreak detection can increase the consistency, scope, and comprehensiveness of detecting hospital-associated transmission.

Introduction

Preventing and containing hospital-associated outbreaks requires timely identification and investigation of possible transmission events. Current methods used by most hospitals to identify clustering of pathogens rely on manual detection of temporal or spatial clustering of a limited number of pre-specified pathogens, often using arbitrary criteria such as ≥3 patients with new methicillin-resistant *Staphylococcus aureus* (MRSA) nosocomial results on a hospital unit within a 2-week period.[1,2] This approach to outbreak detection is problematic for several reasons: 1) it fails to identify outbreaks occurring in pathogens not under routine surveillance, 2) rule-based thresholds for identifying an outbreak, such as a minimum number of cases in a fixed time, can fail to detect some clinically important outbreaks, and can also yield false positive signals, and 3) current methods rely on subjective judgement to determine whether an outbreak exists.

Ideally, hospital-based outbreak detection would utilize automated statistically-based methods to identify clusters across all pathogens, locations, and services, taking into account antimicrobial susceptibility patterns and adjusting for background rates of occurrence.[3] For instance, the minimum number of new isolates that would constitute a cluster would be substantially larger for an endemic organism like methicillin-sensitive *Staphylococcus aureus* than the minimum number for a lower prevalence organism such as a Gram-negative bacillus with an unusual antimicrobial susceptibility pattern.

The use of space-time and higher dimensional scan statistics has been used to detect geographic and temporal clustering of events.[3–8] SaTScan (www.satscan.org) is a free disease surveillance software containing various spatial and space-time scan statistics.[9] This software is widely used to detect and evaluate geographical disease clusters, including applications to detect community disease outbreaks.[10] We integrated SaTScan with WHONET, [11] software that was developed for management and descriptive analysis of microbiology data.[3] It is available from the World Health Organization (WHO) Collaborating Centre for Surveillance of Antimicrobial Resistance. WHONET includes a data-conversion utility (BacLink) that standardizes and imports data from many different microbiology systems into WHONET format. Focusing on microbiology based outcomes in a healthcare microenvironment provides an optimal setting for a statistically-based cluster detection program.

Previous work that applied integrated WHONET-SaTScan software to five years of daily microbiology laboratory data from one 750-bed academic medical center identified an

average of 12 clusters annually.[3] All were deemed to be of clinical interest by hospital epidemiologist reviewers and one-third would have warranted investigation or active intervention had the alerts occurred in real time. In this study, we retrospectively applied this automated outbreak detection system to hospital microbiology data from 43 hospitals to identify statistically-significant clustering of pathogens and asked participating hospitals' infection preventionists and hospital epidemiologists to assess the relevance of each cluster from an operational infection prevention perspective.

Methods

Forty-three hospitals using Premier SafetySurveillor software, an infection prevention surveillance system, participated in this retrospective study, which utilized microbiology and census location data streams routinely sent to Premier for infection prevention surveillance and reporting. Microbiology data were limited to finalized results from specimens obtained greater than 2 days after admission in order to focus on hospital-acquired infections. Hospital microbiology results were processed through WHONET software. We searched for potential clusters of a broad set of microorganisms that have been associated with clusters of hospital-acquired pathogens (Table 1). Only first isolates of a specific organism per hospitalized patient were included and surveillance screens were excluded, as screening is often performed for routine purposes on one day of the week in limited units and practices may change over time leading to false positive clusters. Patient-specific data (patient identifier, admit date) were linked to microbiology data (pathogen name, date of collection). Unit and service two days prior to specimen collection were also collected to identify the location where the organism was likely acquired, consistent with CDC guidance.[12]

Mimicking daily real-time surveillance, we retrospectively applied WHONET-SaTScan to 2 to 3 years of available admission, discharge and transfer and microbiology data accessible through each hospital's Premier SafetySurveillor system data repository. We performed a simulated prospective assessment by adding retrospective data day-by-day to identify when a cluster signaled or progressed. We used WHONET-SaTScan to identify statistically-based clusters, taking into account background rates at each hospital and accounting for multiple testing. We identified clusters of organisms based on hospitals' microbiology results, where clusters were defined by statistically significant increases in pathogens that shared at least one of the following: antimicrobial resistance profile, unit, meta-unit, specialty service, meta-service and whole inpatient hospital. We defined meta-units and meta-services as hospital units and services that commonly share staff and patients such that pathogens could be transmitted between them; for example, a meta-unit might include a cardiology critical care unit and a cardiac surgery intensive care unit.

WHONET-SaTScan parameters were based upon prior studies.[3,13] The maximum length over which a group of isolates could contribute to a detected cluster was set at 60 days; however, there was no restriction on the duration of a cluster if cases continued to accrue with temporally overlapping clusters. Statistical significance was measured using a recurrence interval, which estimated the likelihood that the cluster signal would occur by chance.[13] We used a threshold recurrence interval of 365 days, meaning that a cluster of this type of organism with the observed number and distribution of cases would be expected

to occur by chance less than once per year. Clusters were also restricted to have at least 3 patients.

Infection prevention staff at each hospital were asked to review and assess the value of knowing about cluster signals from WHONET-SaTScan in simulated "real time" using a standardized online survey tool that reflected alerts, if any, day-by-day. The alerts included a summary of the cluster including organism, duration, ward, service, recurrence interval and a line list of the specimens including date, specimen type and resistance pattern when available. They were asked if they would have wanted to have been notified about the cluster, their level of concern and how they would intervene. They were also asked if they had previously been aware of the cluster. If WHONET-SaTScan signaled a second time, meaning that the cluster had expanded, accruing another case, they answered another set of similar questions about the cluster. Finally, they were asked how this automated outbreak detection tool might affect outbreak detection and response at their institution.

We used multivariable logistic regression analysis to assess the association between cluster characteristics and the level of concern about the cluster (moderate or high concern versus low or no concern). Variables included in the model were cluster size (3–4 cases, ≥5 cases); statistical significance, measured by the recurrence interval (≤3 years, >3 years); commensal organisms (yes, no); and pathogen type. Pathogen type was divided into 1) *Clostridioides difficile*, 2) multidrug-resistant organisms (MRSA, vancomycin-resistant *Enterococcus*, and multidrug-resistant Gram-negative rods), 3) and all other pathogens. Analyses were performed with SAS version 9.4 (SAS Institute Inc., Cary, NC).

In addition, a convenience sample of hospitals also provided a list and description of clusters that were previously detected using those hospitals' routine cluster definitions and surveillance methods. Routinely detected clusters were compared to the clusters identified through WHONET-SaTScan. Many of the hospitals did not have complete records of clusters that had been identified and investigated.

Results

Forty-three hospitals participated in the study; these included both academic and community hospitals with a mean bed size of 286 (range 25–913) (Table 2). The hospitals included 42 acute care and one long term acute care facility, and were located in 9 U.S. states.

Using WHONET-SaTScan, we detected a total of 230 clusters involving a broad range of pathogens with an average of 1 cluster per 100 beds per year (range 0–4). Of the 230 clusters, 100 consisted of Gram-positive bacteria, 106 consisted of Gram-negative bacteria, and 24 were clusters of fungi. The pathogens associated with identified clusters are summarized in Table 3 and cluster characteristics are found in Table 4. Clusters were most commonly due to *Enterococcus* sp., *Staphylococcus aureus*, *Escherichia coli*, and *Pseudomonas aeruginosa*. Most clusters consisted of 3–5 patients, but 26% (59) were larger clusters. Approximately half (112) of the clusters had a recurrence interval ranging between 1 and 3 years, and the remaining half were even more statistically uncommon.

Infection prevention programs would have liked to have been notified about 187 (81%) of the clusters and were moderately or highly concerned about 107 (47%) of the clusters (Table 5). They reported being aware of only 24 (10%) of the clusters based upon routine surveillance.

In the multivariate logistic regression analysis, factors associated with whether a cluster was considered moderately or highly concerning included greater statistical significance (recurrence interval >3 years (OR, 1.87; 95% CI 1.01, 3.44)), a large cluster (OR, 4.82; CI 2.47, 9.41), and a cluster of *Clostridioides difficile* or multidrug-resistant organisms (OR, 21.71, 95% CI 2.67, 176.39 and OR, 2.88; CI 1.29, 6.43 respectively) (Table 6). A cluster involving common skin commensal organisms (e.g., *Staphylococcus epidermidis*) was categorized as not important or of low importance (OR 0.19; CI 0.07, 0.55).

Nearly half of the clusters (114) accumulated additional cases after the initial WHONET-SaTScan alert. Of these clusters, 58% (66) were considered to be of moderate to high concern and were assessed as warranting intervention. Of these 66, 41 (62%) involved a pathogen that was not routinely targeted by the hospital's surveillance program. Had the notifications come in real time and the infection preventionists intervened as they indicated, 237 out of 559 cases could potentially have been avoided if interventions successfully prevented further transmission.

In the 10 hospitals that provided a list of clusters identified through routine surveillance, WHONET-SaTScan found that only 9 of 51 (18%) of routinely-detected clusters had a recurrence interval of a year or greater. In addition, WHONET-SaTScan detected an additional 70 clusters that were not previously identified by those hospitals, including 58 (83%) involving pathogens not generally targeted by routine surveillance at those hospitals (e.g., methicillin-sensitive *Staphylococcus aureus*). Among these 10 hospitals, 36 (46%) of the clusters were considered to be of moderate or high concern.

Infection prevention programs from 26 of the 29 hospitals with at least one cluster detected by the automated cluster detection system, noted that they would value the ability to expand the focus of surveillance beyond just multidrug-resistant organisms. All noted the potential for WHONET-SaTScan to improve outbreak detection (22 (76%) to a moderate or large extent) and to streamline their work (18 (62%) to a moderate or large extent).

Discussion

Application of WHONET-SaTScan enhanced the ability of hospitals to detect clusters of hospital-acquired microorganisms by expanding surveillance to include a broad range of pathogens and by automatically assessing for statistically significant clustering of these pathogens by unit, service, and resistance pattern compared to each hospital's baseline data. This outbreak detection tool enabled efficient daily assessments for potential clusters involving 31 pathogens. The tool also provided statistically-derived alerts (e.g. potential outbreaks) that were considered to be of epidemiological importance by infection preventionists at the designed frequency (1 cluster per 100 beds per year).

Application of WHONET-SaTScan in a geographically diverse group of community and teaching hospitals demonstrated that a large number of statistically significant clusters occur across a variety of units, services, and pathogens that are missed by current methods that infection prevention programs use to detect possible outbreaks. An automated outbreak detection tool could allow infection preventionists to identify clusters of pathogens that might otherwise be difficult to recognize because they involve organisms that are not routinely targeted for surveillance (e.g., methicillin-sensitive *Staphylococcus aureus*). Because nearly all hospitals limit their surveillance to focus on MRSA, vancomycin-resistant *Enterococcus*, Gram-negative extended-spectrum beta-lactamase producers, carbapenem-resistant Enterobacteriaceae, and *Clostridioides difficile* for possible outbreaks,[1] it is not surprising that clusters due to other pathogens were missed. We found that more than half of the clusters included organisms not under routine targeted surveillance, and infection preventionists reported that they would have wanted to be notified about the majority of these clusters. In addition, infection preventionists were significantly more likely to judge clusters involving drug-resistant organisms and *Clostridioides difficile* to be moderately or highly concerning, suggesting that infection prevention programs may be less likely to respond to clusters involving other types of pathogens. This tool would greatly expand the outbreak detection capabilities of many hospitals but would require additional education regarding the potential for clusters of relatively antimicrobial-susceptible organisms to also cause hospital-associated outbreaks that increase morbidity and costs.

A large proportion of clusters previously identified through hospitals' routine surveillance were not found to be statistically unusual compared to those hospitals' baseline data. Statistically-based cluster detection would allow focusing of efforts on evaluating and intervening on clusters that are more likely to be clinically significant. Early detection of clusters could enable more rapid implementation of interventions to prevent expansion of hospital-based outbreaks. In this study, approximately 50% of the clusters continued to accumulate cases after the initial statistical alert, suggesting the potential for early interventions triggered by the first cluster signals to prevent subsequent transmission events.

Limitations of the current study include the retrospective nature of these analyses. Although results were provided in simulated real-time, survey responses by infection prevention staff may have been different had the assessments been made prospectively. In addition, other than the 10 hospitals that kept records of work ups for potential outbreaks, many infection prevention programs did not recall or have records about prior clusters for comparison purposes. When WHONET-SaTScan identified a cluster that was also found by routine surveillance, we were often unable to determine the relative timing of the detection due to lack of precise records. Nevertheless, WHONET-SaTScan identified many outbreaks that were not found by routine surveillance.

Automated, statistically-based outbreak detection has the potential to increase the consistency, scope, and comprehensiveness of hospital-associated cluster detection, including clusters due to pathogens not routinely targeted by surveillance. This publicly available tool could be used to trigger rapid responses by infection prevention staff, leading to earlier containment of hospital outbreaks.

Acknowledgements:

We thank Premier, Inc., ChristianaCare, Fairview Health Services, HealthEast Care System, McLeod Health, Meridian Health, Mountain States Health Alliance, Summa Health System, Texas Health Resources, University of California, Irvine, and University Hospitals for their collaboration in this research.

Financial Support

This work was supported by the Centers for Disease Control and Prevention (CDC) Epicenters Program 1U54 CK000172–01 and the National Institutes of Health (NIH) RR025040 and U01CA207167.

References

1. Baker MA, Huang SS, Letourneau AR, et al. Lack of comprehensive outbreak detection in hospitals. Infection Control and Hospital Epidemiology 2016;37:466–8. [PubMed: 26996060]
2. Mellmann A, Friedrich AW, Rosenkotter N, et al. Automated DNA sequence-based early warning system for the detection of methicillin-resistant Staphylococcus aureus outbreaks. PLoS Medicine 2006;3:e33. [PubMed: 16396609]
3. Huang SS, Yokoe DS, Stelling J, et al. Automated detection of infectious disease outbreaks in hospitals: a retrospective cohort study. PLoS Medicine 2010;7:e1000238. [PubMed: 20186274]
4. Kulldorff M A spatial scan statistic. Communications in Statistics - Theory and Methods 1997;26:1481–96.
5. Kulldorff M, Athas WF, Feurer EJ, Miller BA, Key CR. Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. American Journal of Public Health 1998;88:1377–80. [PubMed: 9736881]
6. Kulldorff M Prospective time periodic geographical disease surveillance using a scan statistic. Journal of the Royal Statistical Society: Series A (Statistics in Society) 2001;164:61–72.
7. Kulldorff M, Heffernan R, Hartman J, Assunção R, Mostashari F. A space–time permutation scan statistic for disease outbreak detection. PLoS Medicine 2005;2:e59. [PubMed: 15719066]
8. Natale A, Stelling J, Meledandri M, Messenger LA, D'Ancona F. Use of WHONET-SaTScan system for simulated real-time detection of antimicrobial resistance clusters in a hospital in Italy, 2012 to 2014. Euro Surveill 2017; 22(11):30484. [PubMed: 28333615]
9. Kulldorff M and IMS Inc. SaTScan v9: Software for the spatial, temporal and space-time scan statistics; www.satscan.org Accessed Nov 19, 2019.
10. Green SK, Peterson ER, Kapell D, Fine AD, Kulldorff M. Daily reportable disease spatiotemporal cluster detection, New York City, New York, USA 2014–2015. Emerg Infect Dis 2016; 22(10):1808–1812. [PubMed: 27648777]
11. Stelling JM, O'Brien TF. Surveillance of antimicrobial resistance: the WHONET program. Clinical Infectious Diseases: an official publication of the Infectious Diseases Society of America 1997;24 Suppl 1:S157–68. [PubMed: 8994799]
12. Identifying Healthcare-associated Infections (HAI) for NHSN Surveillance; https://www.cdc.gov/nhsn/pdfs/pscmanual/2psc_identifyinghais_nhsncurrent.pdf Accessed Nov 10, 2019.
13. Kleinman K, Lazarus R, Platt R. A generalized linear mixed models approach for detecting incident clusters of disease in small areas, with an application to biological terrorism. American Journal of Epidemiology 2004;159:217–24. [PubMed: 14742279]

Table 1.

Priority pathogens

Pathogen
Absidia sp.
Achromobacter sp.
Acinetobacter sp.
Aeromonas sp.
Alcaligenes sp.
Aspergillus sp.
Bacteroides sp.
Burkholderia sp.
Candida sp.
Citrobacter sp.
Clostridioides difficile *
Cunninghamella sp.
Cutibacterium sp.†
Enterobacter sp. ‡
Enterococcus sp.
Escherichia sp.
Fusarium sp.
Group A Streptococcus
Haemophilus sp.
Klebsiella sp. ‡
Legionella sp.
Malassezia sp.
Mycobacterium sp.
Proteus sp.
Pseudomonas sp.
Raoultella sp.
Rhizopus sp.
Salmonella sp.
Serratia sp.
S. aureus
Stenotrophomonas sp.

*Previously *Clostridium*

†Previously *Propionibacterium*

‡*Klebsiella aerogenes* included with *Enterobacter sp.*, previously classified as *Enterobacter aerogenes*

Table 2.

Characteristics of the participating hospitals

Health Care Facility*	# of participating health care facilities (N=43)	Mean bed size
Community		
<200 beds	21	90
≥200 beds	7	302
Teaching		
<200 beds	0	
≥200 beds	14	582

* 1 Long term acute care facility, not included due to potential identifiability

Table 3.

Clusters

Pathogen	Number of clusters N (%)
Enterococcus sp. (including VRE)	37 (16)
*Staphylococcus aureus (*including MRSA*)*	32 (14)
Escherichia coli	29 (13)
Pseudomonas aeruginosa	28 (12)
Klebsiella sp.*	23 (10)
Candida sp	21 (9)
Clostridioides difficile †	17 (7)
Enterobacter sp.*	11 (5)
Staphylococcus sp.	11 (5)
Serratia sp.	7 (3)
Stenotrophomonas sp	4 (2)
Aspergillus sp.	3 (1)
Cutibacterium sp. ‡	2 (1)
Proteus sp.	2 (1)
Acinetobacter sp	1 (<1)
Burkholderia sp.	1 (<1)
Haemophilus sp.	1 (<1)
Total	230

*Klebsiella aerogenes included with Enterobacter sp., previously classified as Enterobacter aerogenes

†Previously Clostridium

‡Previosssusly Propionibacterium

Table 4.

Characteristics of clusters

Cluster Characteristics	N (%)	Clusters that progressed (%)	Mean # Clusters per Hospital (Range)	Clusters/100 beds/year Mean (Range)
Total number of clusters	230 (100%)	114 (50%)	5.35 (0–41)	1.02 (0–3.65)
Organisms				
Gram-positive	100 (44%)	52 (46%)	2.33 (0–17)	0.45 (0–1.82)
Gram-negative	106(46%)	46 (40%)	2.47 (0–18)	0.47 (0–1.62
Fungi	24 (10%)	16 (14%)	0.56 (0–5)	0.11 (0–1.52)
Alert type				
Resistance profile	104 (45%)	52 (46%)	2.42 (0–14)	0.46 (0–1.58)
Unit	86 (37%)	35 (31%)	2.0 (0–6)	0.38 (0–1.46)
Service	99 (43%)	62 (54%)	2.30 (0–19)	0.44 (0–1.52)
Hospital-wide	27 (12)	17 (15%)	0.63 (0–3)	0.12 (0–0.91)
Size				
3–5	171 (74%)	67 (59%)	3.98 (0–30)	0.75 (0–3.64)
6–10	38 (17%)	28 (25%)	0.88 (0–10)	0.17 (0–0.91)
>10	21 (9%)	19 (17%)	0.49 (0–3)	0.09 (0–0.45)
Recurrence interval				
1–3 years	112 (49%)	24 (21%)	2.60 (0–15)	0.49 (0–2.55)
>3–10 years	48 (21%)	28 (25%)	1.12 (0–14)	0.21 (0–1.09)
>10 years	70 (30%)	62 (54%)	1.63 (0–12)	0.31 (0–1.52)

Table 5.

Assessment of clusters

Question	Initial signal N (%)	Follow-up signal N (%)
Desire notification of cluster		
Yes	162 (70%)	97 (85%)
No	68 (30%)	17 (15%)
Level of concern		
No concern	64 (28%)	18 (16%)
Low concern	97 (42%)	33 (29%)
Moderate or high concern	69 (30%)	63 (55%)
Action		
Notify other members of IP program	50 (22%)	49 (43%)
Assess background frequency	38 (17%)	33 (29%)
Notify manager of unit/service	42 (18%)	37 (33%)
Assess medical records for common source characteristics	50 (22%)	50 (44%)
Activate response measures	29 (13%)	29 (25%)
Decision factors (important/very important)		
Organism	221 (96%)	108 (95%)
Source	198 (86%)	98 (86%)
Antibiotic profile	177 (77%)	89 (78%)
Location of the cluster	215 (94%)	104 (91%)
Statistical likelihood of cluster	181 (79%)	98 (85%)

Table 6.

Factors for consideration of cluster as moderately or highly concerning

Variable	OR (95% CI)
RI (>3 years, ≤ 3 years)	1.87 (1.01, 3.44)
Size (≥5 cases, < 5 cases)	4.82 (2.47, 9.41)
Pathogen type	
Not MDRO, not *C. difficile*	1.00
C. difficile	21.71 (2.67, 176.39)
MRSA, VRE, MDR GNR	2.88 (1.29, 6.43)
Commensal organism	0.19 (0.07, 0.55)