ARTICLE IN PRESS

American Journal of Infection Control ■■ (2017) ■■-■■

Contents lists available at ScienceDirect

American Journal of Infection Control

journal homepage: www.ajicjournal.org

Major Article

Methods for computational disease surveillance in infection prevention and control: Statistical process control versus Twitter's anomaly and breakout detection algorithms

Timothy L. Wiemken PhD, MPH, FAPIC, CIC [a,*], Stephen P. Furmanek MPH, MS [b], William A. Mattingly PhD [b], Marc-Oliver Wright MT(ASCP), MS, CIC, FAPIC [c], Annuradha K. Persaud MPH [b], Brian E. Guinn PhD, MPH [b], Ruth M. Carrico PhD, RN, FNP-C, FSHEA, CIC [b], Forest W. Arnold DO, MSc [b], Julio A. Ramirez MD [b]

[a] *Department of Epidemiology and Population Health, School of Public Health and Information Sciences, University of Louisville, Louisville, KY*
[b] *Healthcare Epidemiology and Patient Safety Program, Division of Infectious Diseases, University of Louisville, Louisville, KY*
[c] *Department of Infection Prevention and Control, University of Wisconsin Hospitals and Clinics, Madison, WI*

Key Words:
Health care epidemiology
Patient safety
Outbreak
Multidrug resistant organisms
Device-associated infections

Background: Although not all health care-associated infections (HAIs) are preventable, reducing HAIs through targeted intervention is key to a successful infection prevention program. To identify areas in need of targeted intervention, robust statistical methods must be used when analyzing surveillance data. The objective of this study was to compare and contrast statistical process control (SPC) charts with Twitter's anomaly and breakout detection algorithms.

Methods: SPC and anomaly/breakout detection (ABD) charts were created for vancomycin-resistant *Enterococcus*, *Acinetobacter baumannii*, catheter-associated urinary tract infection, and central line-associated bloodstream infection data.

Results: Both SPC and ABD charts detected similar data points as anomalous/out of control on most charts. The vancomycin-resistant *Enterococcus* ABD chart detected an extra anomalous point that appeared to be higher than the same time period in prior years. Using a small subset of the central line-associated bloodstream infection data, the ABD chart was able to detect anomalies where the SPC chart was not.

Discussion: SPC charts and ABD charts both performed well, although ABD charts appeared to work better in the context of seasonal variation and autocorrelation.

Conclusions: Because they account for common statistical issues in HAI data, ABD charts may be useful for practitioners for analysis of HAI surveillance data.

© 2017 Association for Professionals in Infection Control and Epidemiology, Inc. Published by Elsevier Inc. All rights reserved.

Each year, nearly 750,000 health care-associated infections (HAIs) cause approximately 75,000 excess deaths.[1] Although not all HAIs are preventable, elimination is the goal of most, if not all, infection prevention and control (IPC) programs. To identify significant changes in disease trends, as well as to document the effectiveness of IPC interventions with minimal statistical error, disease surveillance methodologies with robust statistical analyses must be used. Sound statistical methods may allow infection preventionists (IPs) to more confidently identify these factors. Statistical process control (SPC) is a common statistical method for surveillance used for defining an expected baseline disease or process compliance rate, and for monitoring processes for positive or negative abnormal variation.[2,3] This results in identification of areas where improvements are necessary or where interventions have been successful.[4,5] Although many different computer software programs are available to create SPC charts, many are costly due to software licensing and training, and their output can be complex. Other traditional statistical methodologies can be used to evaluate HAI data, but the overwhelming majority of these basic tests (eg, χ^2 tests, *t* tests, logistic/linear regression, and Shewhart SPC charts) may be inappropriate for HAI surveillance due to issues such as process changes,

* Address correspondence to Timothy L. Wiemken, PhD, MPH, FAPIC, CIC, University of Louisville, Department of Epidemiology and Population Health, School of Public Health and Information Sciences, 485 E Gray St #230, Louisville, KY 40202.
E-mail address: tim.wiemken@louisville.edu (T.L Wiemken).
Conflicts of interest: None to report.

0196-6553/© 2017 Association for Professionals in Infection Control and Epidemiology, Inc. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ajic.2017.08.005

ARTICLE IN PRESS

organism seasonality,[6] and autocorrelation (ie, rates this month are more closely related to rates last month than they are rates from last year).[7] Traditional SPC charts suffer from similar statistical issues, and although some nontraditional SPC charts may account for seasonality and autocorrelation, they are rarely used or discussed in IPC, and are not currently implemented in the National Healthcare Safety Network (NHSN). Further, they often require creation of multiple charts, each using several rules for evaluating variation, resulting in difficult interpretation. Novel surveillance methods that may account for the unique aspects of HAI data are needed. Algorithms for detection of anomalous (ie, abnormal) data points and breakouts (ie, shifts in the mean or gradual ramp up/down in data over time) that account for seasonality and autocorrelation, and allow for the visualization of data have recently been open-sourced by Twitter Inc (San Francisco, CA).[8-10] Although Twitter has been used for data collection in the field of infectious diseases surveillance,[11,12] currently, no data exist regarding these new algorithms' potential utility for HAI data surveillance.

The objective of this study was to compare and contrast Twitter's anomaly and breakout detection algorithms with traditional SPC charts for HAI surveillance. Further, we provide access to a free tool built by our team for anyone to create report-ready figures for anomaly/breakout detection and traditional SPC charts.

METHODS

Patients, setting, and study design

This was a secondary analysis of data collected by an IPC department for patients requiring intensive care at a 404-bed, metropolitan, level-1 trauma center in Kentucky.

Inclusion and exclusion criteria

All patients admitted to intensive care units at this facility from January 2009-June 2016 were evaluated for inclusion in the analysis. Because some data were only available for a subset of this time frame, some analyses were of shorter duration.

Study definitions

Microbiologic culture data and device-associated infection data were included in this study. For microbiologic data, health care-associated, health care-onset vancomycin-resistant *Enterococcus* (VRE), and *Acinetobacter baumannii* were evaluated. For microbiologic data, the numerator consisted of the first positive culture for each patient each month. Surveillance cultures and subsequent positive cultures for the same patient during the same month were not included. The denominator consisted of the aggregate number of patient-days each month. For device-associated infections, catheter-associated urinary tract infection (CAUTI) and central line-associated bloodstream infection (CLABSI) data were evaluated, and NHSN definitions were used for surveillance purposes,[13,14] and direct downloads from the NHSN were used for data gathering. To calculate rates of device-associated infections, the numerator consisted of the number of cases each month, whereas the denominator consisted of the aggregate number of device-days each month (urinary catheter-days for CAUTI and central-line days for CLABSI). For CLABSI data, additional charts were created using a 15-point subset of the original data. Although NHSN device-associated infection surveillance definitions changed for CAUTI and CLABSI during the study period, the objective of this study was a comparison of methods and did not warrant further data adjustment.

Human subjects protection

This study was approved by the University of Louisville Human Subjects Protection Program Office (protocol No. 05.0556).

Statistical analysis

For each item under study, both SPC and anomaly/breakout detection (ABD) charts were constructed. For device-associated infection rates, SPC u charts were constructed, whereas p charts were used for microbiologic data. Data points indicating special-cause variation were documented on the charts using large white dots (red color in the Web application) if any of the following criteria were met: 1 point above/below $\pm 3\sigma$, 2 of 3 consecutive points above/below $\pm 2\Sigma$, 4 of 5 consecutive points above/below $\pm 1\Sigma$, and 8 consecutive points above/below the mean. These represent a subset of the original Montgomery rules, often used in the health care setting.[15]

The ABD algorithms were run for each item under study and ABD charts were constructed to display the data. Unexpected spikes in data (ie, anomalies) were detected using Twitter's anomaly detection algorithm.[8,9] This algorithm uses time series decomposition and a seasonal hybrid generalized extreme studentized deviate test for outliers. Breakouts, or segmented shifts in the mean and/or gradual ramp up/down from 1 steady state to another, were detected using Twitter's breakout detection algorithm.[10] This approach uses the E-devisive with medians algorithm to detect divergence in the mean or a change in the data distribution in a time series. The mathematics behind these algorithms are beyond the scope of this study, but in-depth reviews can be found elsewhere in the literature.[8,16]

The ability of these algorithms to decompose the data into seasonal and autocorrelative components may provide a more rigorous evaluation of trends, compared with simply comparing a single data point to an average of the prior data as is commonly done with traditional methods. Anomalies on ABD charts were annotated similar to special causes on the SPC charts, whereas breakouts were documented as horizontal dotted lines. For anomaly detection, a maximum of 5% of the data were allowed to be anomalous. A minimum of 3 consecutive data points were chosen to indicate a breakout by default.

R version 3.3.1 (R Foundation for Statistical Computing, Vienna Austria) was used for all analyses. R version 3.3.1 and Shiny version 1.0.3[17] (R Foundation for Statistical Computing) were used to create a Web site for automated creation of both ABD and SPC charts while providing users control over the several variable inputs for ABD and SPC charts, including maximum percent of anomalies, minimum points for a breakout, standardization factors, y-axis labels, and benchmark rates. This application also warns the user when judgments may be invalid due to statistical issues with the data. Furthermore, all charts created are able to be downloaded through the interface and no data or charts are saved, providing a safe and anonymous environment for report creation on our secure Web server hosted at our university. The Web tool is free to use and is located at https://crsp.louisville.edu/shiny/anomaly. The Web application requires only a date, monthly aggregate numerator (eg, total infections per month) and monthly aggregate denominator (eg, total device-days / patient-days per month), no protected health information is needed.

RESULTS

For each organism and device-associated infection, SPC charts as well as ABD charts were created.

Figure 1 depicts the SPC p chart and ABD chart for VRE. Special-cause variation was identified on the SPC chart in January 2013, a point that was also detected as anomalous on the ABD chart. An

Download English Version:

https://daneshyari.com/en/article/8567020

Download Persian Version:

https://daneshyari.com/article/8567020

Daneshyari.com