BUILDING A SURVEILLANCE SYSTEM FOR BIOTERRORISM AND COMMUNITY OUTBREAKS

The Use of Multiple Algorithms to Detect Unusual Increases in Hospital Admissions, Visits, and Deaths in Los Angeles County

BACKGROUND

Los Angeles city was scheduled to host the Democratic National Convention (DNC) in August 2000. By June 2000, Acute Communicable Disease Control Unit (ACDC) of the Los Angeles County Department of Health Services (LAC DHS) began collaborations with the Centers for Disease Control and Prevention (CDC) and the California State Health Department. By the time the Convention started, a team of approximately ten health professionals was conducting surveillance for certain symptoms suggestive of bioterrorism. Aiming for timeliness, the multi-agency team devoted many resources in terms of time and people.

An ACDC epidemiologist and graduate-level student worker proposed an alternate surveillance system. For simplicity of data collection and analysis, they decided to use daily counts of admissions, visits, and deaths, instead of counts of diagnostic codes for syndromic surveillance. Using established outbreak detection algorithms, the surveillance system would detect statistically significant (p 0.05) increases in the daily counts.

ACDC has conducted previous research on outbreak detection algorithms. Using four outbreak detection algorithms in studying campylobacteriosis, ACDC attempted to validate the algorithms retrospectively (Special Reports 1999). However, a prospective approach may be more appropriate for validation because outbreaks can be separated more easily between those detected by an increase in reports and those detected by a single phone call from a physician or patient describing a possible outbreak situation.

Thus, there were two proposed purposes for establishing this bioterrorism surveillance system. The first purpose was to explore using outbreak detection algorithms modified to analyze daily data instead of weekly data to alert when possible bioterrorist events or community outbreaks occur. The second purpose was to attempt to validate the algorithms prospectively.

METHODS

Data Collection

On July 1, 2000, under the directive of the former Associate Director of Clinical and Medical Affairs Harbor-UCLA Medical Center (Harbor), Martin Luther King/Drew Medical Center (KDMC), Los Angeles County+University of Southern California Medical Center (LAC+USC), and Olive View Medical Center (OVMC) started providing ACDC with daily counts of hospital admissions, intensive care unit (ICU) admissions, emergency department (ED) visits, and total hospital deaths. These four of the six LAC medical centers were chosen because of their larger size and relatively more central location. The four medical centers provided daily counts data from July 1, 1999 to June 30, 2000 to serve as baseline information for the algorithms. ACDC received the data by facsimile or electronic mail according to the hospital data categories in Table 1.

Table 1. Categories of Data by Medical Center

Categories of data by medical center*			
LAC+USC	KDMC	Harbor	OVMC
Total hospital	Total hospital	Total hospital	Total hospital
Total ICU	Total ICU	Total ICU	Total ICU
Total deaths	Total deaths	Total deaths	Total deaths
General ED	Acute ED	Adult ED	
OB-GYN ED	Pediatric ED	Pediatric ED	
Pediatric ED	Psychiatric ED	Psychiatric ED	
Psychiatric ED			

* LAC+USC is Los Angeles County+University of Southern California Medical Center, KDMC is Martin Luther King/Drew Medical Center, Harbor is Harbor/UCLA Medical Center, and OVMC is Olive View Medical Center.

Data Management

Using Microsoft Excel, ACDC developed a database that would organize the data and automatically perform statistical calculations and provide graphical output. The data were transferred or manually entered into the database on a daily basis. In order to quickly focus on Total hospital, Total ICU, Total deaths, and Total ED data, the category of Total ED was made by combining the counts of all EDs.

The Algorithms

The outbreak detection algorithms included those used by the CDC in the *Morbidity and Mortality Weekly Report (MMWR)*, the Current Day (CD) method used by the World Health Organization to identify meningococcal meningitis outbreaks, the Current-Previous method used by the Oregon State Health Department, and the CuSum method, also known as the Salmonella Outbreak Detection Algorithm (SODA) used by CDC. These algorithms were modified to analyze daily data instead of weekly data and to explore different levels of "sensitivity" by varying the lengths of reference periods.

Figure 1. Calculation of MMWR algorithm.

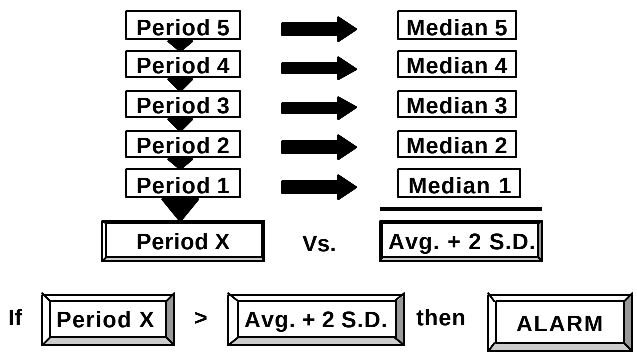
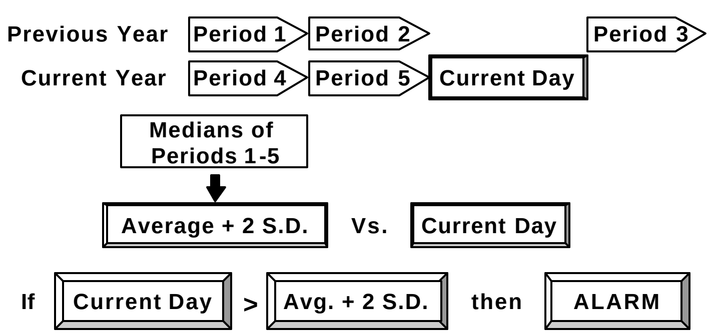


Figure 2. Calculation of Current-Previous algorithm.



The *MMWR* method used the medians of the counts for five previous periods centered on days of the same name as the current day of interest to find an average (Figure 1). This average plus two standard deviations served as the reference value to which the count for the day of interest was compared. The algorithm triggered an alarm when the count for the day of interest exceeded the reference value. The authors developed two versions of this algorithm. The first version, designated MMWR3, referenced a three-day time period, and the second, designated MMWR7, referenced a seven-day time period. For example, in applying MMWR3 to a Monday (Period X in Figure 1), the medians of the daily counts of the five previous three-day sets of Sunday, Monday, and Tuesday would be used to calculate an average. The addition of two standard deviations to the average would define a reference value that would be compared to the count for the current Monday. The investigators modified this algorithm to account for certain weekdays being busier than others in terms of admissions and visits.

The Current-Previous algorithm averaged the medians of the daily counts of five time periods (Figure 2). To account for seasonality, the first three time periods were of the previous year. The two versions of this algorithm were designated CP3 and CP7. CP3 referenced a time period of three days and CP7 referenced a time period of seven days. For example (Figure 2), using the three-day version of this algorithm, the authors would define Period 5 as the three days before the day of interest, Period 4 as the three days before Period 5, Period 1 as the same calendar days as Period 4 but of the previous year, Period 2 as the same calendar days of Period 5 but of the previous year, and Period 3 as the three calendar days after the current day of interest but of the previous year. The medians of the daily counts of these five three-day periods would be averaged. The average plus two standard deviations made the reference value which would be compared to the count for the day of interest.

When the count for the day of interest exceeded the reference value, the algorithm triggered an alarm.

The Current-Day (CD) method compared an average of the daily counts of the previous four days plus two standard deviations to the count of the day of interest. When the count for the day of interest was greater, the algorithm triggered an alarm (Figure 3).

Figure 3. Calculation of Current Day algorithm.

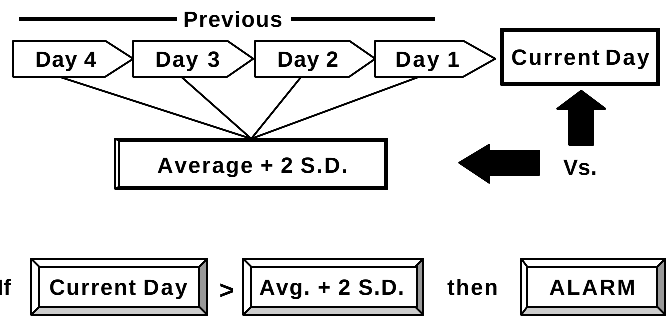
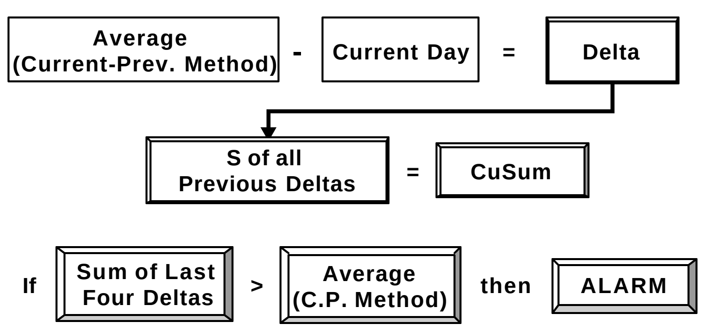


Figure 4. Calculation of CuSum algorithm.



There were three-day, seven-day, and fourteen-day versions of the CuSums method. The reference value of the CuSums method was calculated exactly like the reference value in the Current-Previous method but did not include the addition of two standard deviations. The difference between the CuSum reference value and the count for the day of interest provided a value designated *delta*. The cumulative sum of the *deltas* was the CuSum. The sum of the deltas of the day of interest and the previous three days was compared to the reference value. When the sum of the four deltas exceeded the reference value, the algorithm triggered an alarm (Figure 4). A line graph of the CuSums provided a sensitive visual tool of trends as positive slopes indicated increases in admissions, visits, and deaths. Alarms from the CuSum algorithms were considered as secondary alarms because they tended to alarm or not alarm for extended periods of time. In other words, the CuSum alarms tended to cluster.

Surveillance

The first step of looking at the alarms involved a daily summary of the total alarms in each medical center. For example, if the current date is December 12, 2000, the number of total alarms and the number of alarms within each hospital data category would be reviewed as shown in Table 2. The medical centers with any alarms would receive closer review. In the example, LAC-USC was the only medical center with alarms.

Table 2. Example of Daily Summary of Alarms by Medical Center

Current date: 12/12/2000

		Number of alarms by hospital data category			
Medical center*	Total Alarms	Hospital admissions	ICU admissions	Total deaths	Total ED
Harbor	0	0	0	0	0
LAC-USC	12	0	6	4	2
KDMC	0	0	0	0	0
OVMC	0	0	0	0	N/A

* LAC+USC is Los Angeles County-University of Southern California Medical Center, KDMC is Martin Luther King/Drew Medical Center, Harbor is Harbor/UCLA Medical Center, and OVMC is Olive View Medical Center.

The next step involved focusing on the medical centers with alarms and within each hospital data category comparing the number of previous alarms to the current number. To continue the example of December 12, 2000, Table 3 focuses on LAC+USC and shows that in the past two days there have been 11 and 12 total alarms but unlike the current day the majority of these occurred in hospital data categories other than the ICU. Since the majority of the alarms on December 12 occurred in the ICU, the next step would involve looking at this hospital data category.

Table 3. Example of Summary of One Medical Center by Hospital Data Category

Medical center: LAC-USC

Date in 2000	Day of the week	Total Alarms	Number of alarms by hospital data category: Hospital admissions	ICU admissions	Total deaths	Total ED
12/10	Sun	11	0	3	7	1
12/11	Mon	12	0	0	6	6
12/12	Tue	12	0	6	4	2
12/13	Wed	6	2	1	1	2
12/14	Thu	9	0	4	0	5
12/15	Fri	7	0	3	0	4

Table 4. Example of Summary of Hospital Data Category by Algorithm

Hospital data category: ICU   Medical center: LAC-USC

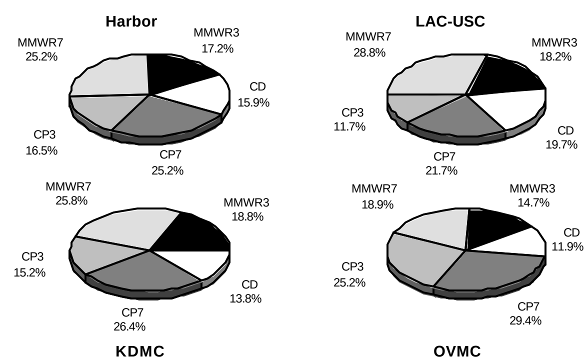
Date in 2000	Algorithms that alarmed: MMWR 3	MMWR 7	Current-Preveious 3	Current-Preveious 7	Current Day	CuSums 3	CuSums 7	CuSums 14	Total alarms
12/10						X	X	X	3
12/11									0
12/12	X	X		X		X	X	X	6
12/13							X		1
12/14	X	X		X			X		4
12/15						X	X	X	3

Seeing which algorithms caused certain hospital data categories to have relatively more alarms would be the next step in surveillance. Each algorithm refers to a particular time period and so several algorithms triggering alarms should warn the investigators that an unusual increase has occurred. Continuing the example of December 12, 2000, Table 4 shows MMWR3, MMWR7, and CP7 triggering alarms for the first time since two days earlier. The total number of alarms for the previous seven days, month, and three months would be graphed with a vertical bar chart (data not shown) to give the investigators more perspective on the current day's alarms.

RESULTS

LAC-USC continued to provide daily count data into the year 2001; however, the other medical centers did not. OVMC stopped providing data on September 6, Harbor stopped on November 30, and KDMC stopped on December 4.

Figure 5. Percent of Alarms by Algorithm, 7/1/00-9/6/00.



From July 1, 2000 to September 6, 2000, MMWR7 and CP7 usually triggered the most alarms (Figure 5). Regarding Total hospital (hospital admissions) data, OVMC experienced the greatest number of alarms, 38, while Harbor experienced 26 alarms, and LAC-USC and KDMC each had 12. Regarding ICU data, OVMC experienced 80 alarms, while LAC-USC experienced 44, Harbor 35, and KDMC 26. Regarding Total deaths data, Harbor experienced 49 alarms, while KDMC experienced 34, LAC-USC 31, and OVMC 25. Regarding Total ED visits, Harbor experienced 26 alarms, LAC-USC 22, and KDMC 21 (OVMC did not have this hospital data category).

DISCUSSION

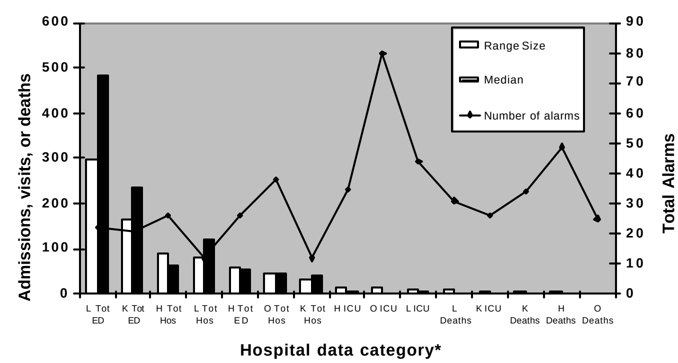
The intended purpose of this surveillance system was to detect unusual increases in hospital admissions, ED visits, ICU admissions, and hospital deaths that might indicate an unsuspecting population's exposure to a bioterrorist weapon or a naturally occurring infectious disease. Unfortunately, the surveillance system lost participants after the DNC ended and after the Associate Director of Clinical and Medical Affairs left his LAC position. One of the misconceptions that might have contributed to the drop in participation was that the threat of bioterrorism is only real during large national or international gatherings like the DNC or the Olympics. A surveillance system with set protocols for easy and efficient maintenance can only be established when government officials (local, state, and national) realize that bioterrorism has a potential to occur at any time and affect large populations for an extended period.

Although one of the purposes of this study was to validate the modified algorithms, the lack of any protocols to do this left the questions of predictive value positive and negative unanswered. In other words, how often the algorithms alarmed when outbreaks of community illness occurred and how often they did not alarm when possibilities of common-source community illness were absent were not measured. Certain algorithms triggered more alarms than others. However, this should not be interpreted that these algorithms are more sensitive and more useful than others. The investigators used several algorithms because the alarms might indicate not just outbreaks, but various unrelated illnesses, hospital-specific practices, random variation, or low baseline data. One method of validation could involve collecting diagnostic or admission data with chief complaints. However, any method of validation most probably would require hospital staff to collect the data.

A related limitation of the surveillance system was the lack of protocols for investigation at the medical center. When algorithms alarmed, there were no established procedures that stated who would investigate and what the investigation would entail. Access to diagnostic data or chief complaint data at admission could help assess the presence of any common exposures or risk factors.

Part of the problem in establishing investigation protocols was the one-day delay (three-day delay for Mondays) in receiving the data. The earliest the alarms could trigger was one day after the medical centers attended to their patients. In the case of bioterrorist events, this delay might be crucial to saving many lives. Furthermore, the one-day delay often increased as most of the medical centers did not have a consolidated database and collected the daily counts by hand tallies in some instances. Thus, the busyness of one department may delay the daily count data from being transmitted. Indeed, the method of facsimile and E-mail dictated a data management system requiring a large amount of manual entry and daily updating of data analysis programs at ACDC. The necessity of such tasks hurt the authors' ability to establish and coordinate any investigative efforts at the medical centers.

Figure 6. Ranges and Medians of Daily Counts by Number of Alarms, 7/1/00-9/6/00.



* L =LAC -USC, K=KDMC, H=Harbor, O=OVMC

The size of the medical center and its facilities for the various categories of hospital data could have played a role in the number of alarms that triggered. As one might have expected, hospital data categories with low numbers of daily counts and small ranges of values demonstrated a greater potential for alarms to trigger (Figure 6). For example, from July 1, 2000 to September 6, 2000, total hospital deaths in Harbor had a median daily count of one, a range of zero to five deaths in a day, and the second largest number of alarms among all hospital data categories (49 alarms). Some of these hospital data categories had periods where a majority of the algorithms triggered alarms. Investigation of alarms might help determine thresholds and guidelines that would increase the predictive values of the surveillance system.

With respect to the surveillance conducted by the collaboration of ACDC, CDC, and the California State Health Department, the surveillance conducted here was established and maintained by one epidemiologist and one student worker. Thus, this surveillance system is relatively cost-effective. Although the modifications of established disease outbreak algorithms seem to work with the study data, the question of validity remains. Answering this question requires close collaborations with medical centers.

RECOMMENDATION

To evaluate the usefulness of this surveillance system and improve it, a formal relationship between ACDC and any participating medical center needs to be established such that standard protocols are defined and followed for investigating causes of alarms. ACDC and medical centers need to collaborate in developing databases that allow for real-time or near real-time transmission of data. Again, the lack of centralized databases in medical centers contributes to the delay in ACDC getting information. In addition, chief complaint and diagnostic data at time of observation of admission, visit, and death could only help to quickly evaluate the algorithms. This evaluation could suggest the need to modify the algorithms further to improve predictive values.