

# Use of the CUSUM and Shewhart control chart methods to identify changes of public health significance using childhood blood lead surveillance data

Timothy Dignam<sup>a</sup>, James Hodge<sup>b</sup>, Stella Chuke<sup>b</sup>, Carlos Mercado<sup>c</sup>, Adrienne S. Ettinger<sup>b</sup>, W. Dana Flanders<sup>a</sup>

**Background:** Local, state, and national childhood blood lead surveillance is based on healthcare providers and clinical laboratories reporting test results to public health departments. Increased interest in detecting blood lead level (BLL) patterns and changes of potential public health significance in a timely manner has highlighted the need for surveillance systems to rapidly detect and investigate these events.

**Objective:** Decrease the time to detect changes in surveillance patterns by using an alerting algorithm developed and assessed through historical child blood lead surveillance data analysis.

**Methods:** We applied geographic and temporal data-aggregation strategies on childhood blood lead surveillance data and developed a novel alerting algorithm. The alerting algorithm employed a modified cumulative summary/Shewhart algorithm, initially applied on 113 months of data from two jurisdictions with a known increase in the proportion of children <6 years of age with BLLs  $\geq 5 \mu\text{g/dl}$ .

**Results:** Alert signals retrospectively identified time periods in two jurisdictions where a known change in the proportion of children <6 years of age with BLLs  $\geq 5 \mu\text{g/dl}$  occurred. Additionally, we identified alert signals among six of the 18 (33%) randomly selected counties assessed where no previously known or suspected pattern changes existed.

**Conclusion:** The modified cumulative summary/Shewhart algorithm provides a framework for enhanced blood lead surveillance by identifying changes in the proportion of children with BLLs  $\geq 5 \mu\text{g/dl}$ . The algorithm has the potential to alert public health officials to changes requiring further important public health investigation.

**Key Words:** surveillance, blood lead level, child, alerting algorithm, cumulative summary (CUSUM)/Shewhart control charts

## Introduction

The Flint, Michigan water crisis, detected in part by an astute pediatrician,<sup>1</sup> highlighted the need for enhanced surveillance of

blood lead levels (BLLs) and public health response.<sup>2</sup> Improved capacity for state and local health departments to frequently analyze trends in reported blood lead test results (e.g., monthly) can potentially decrease the time required to detect deviations of public health significance. Rapid analysis of surveillance data to detect unusual patterns in BLLs is an early indicator for intervention at the local level. Childhood blood lead surveillance (CBLS) data provide a valuable opportunity to adapt time-series analytic methods for this purpose. Cumulative summary (CUSUM) and Shewhart control charts compare differences between observed and expected results using time series data in the context of a threshold. The CUSUM chart is the normalized residuals beyond a reference interval that is cumulatively summed sequentially. A Shewhart chart checks the residual compared to an expected value of the moving range of the residuals. An alert is generated when a plotted value exceeds a threshold value. Our approach was to use cumulative sums

<sup>a</sup>Division of Environmental Health Science and Practice, National Center for Environmental Health, Centers for Disease Control and Prevention, Atlanta, Georgia; <sup>b</sup>Lead Poisoning Prevention and Environmental Health Tracking Branch, Division of Environmental Health Science and Practice, National Center for Environmental Health, Centers for Disease Control and Prevention, Atlanta, Georgia; and <sup>c</sup>Guidehouse LLP, Washington, District of Columbia.

**SDC** Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article ([www.enviroepidem.com](http://www.enviroepidem.com)).

The authors declare that they have no conflicts of interest with regard to the content of this report.

The findings, conclusions, and views expressed in this manuscript are those of the authors and do not necessarily represent the views, policies or official position of the US Centers for Disease Control and Prevention.

Limited, de-identified child blood lead surveillance data are available from the CDC.

\*Corresponding Author: Division of Environmental Health Science and Practice, National Center for Environmental Health, Centers for Disease Control and Prevention, Atlanta, GA 30341. E-mail: [tdignam@cdc.gov](mailto:tdignam@cdc.gov) (T. Dignam).

Written work prepared by employees of the Federal Government as part of their official duties is, under the U.S. Copyright Act, a "work of the United States Government" for which copyright protection under Title 17 of the United States Code is not available. As such, copyright does not extend to the contributions of employees of the Federal Government.

*Environmental Epidemiology* (2020) 4:e090

Received: 27 November 2019; Accepted 14 February 2020

Published online 31 March 2020

DOI: 10.1097/EE9.000000000000090

## What this study adds

This manuscript describes a new, validated alerting algorithm to rapidly analyze child blood lead surveillance data and alert health department authorities to potential spikes of elevated blood lead levels requiring public health investigation. There is no such syndromic surveillance method currently applied to child blood lead surveillance in the United States or other countries. State surveillance databases are potentially underutilized—mostly used to create periodic reports—while also a critical tool to identify increases in child lead exposure. Application of this algorithm has the potential to enhance the child lead poisoning prevention surveillance landscape—to potentially mirror infectious disease syndromic surveillance. The article describes the successful evaluation of the method on data from 20 US counties/jurisdictions.

of observed values minus expected values, where the expected values are adjusted for covariates.<sup>3</sup>

Approximately 500,000 US children have blood lead test results higher than the Centers for Disease Control and Prevention's (CDC) blood lead reference level (BLL  $\geq 5$   $\mu\text{g}/\text{dl}$ ). This reference level represents the 97.5th percentile of BLL distribution among children less than 6 years of age.<sup>4</sup> CDC funds childhood lead poisoning prevention programs in state and local health departments through cooperative agreements to conduct primary and secondary prevention of child lead exposure. These programs receive BLL test results continuously from healthcare providers and clinical laboratories as required by state laws or regulations. Most states require clinical laboratories and healthcare providers to report all blood lead test results to state health departments. Since 1997, funded programs have submitted CBLs data to CDC based on cooperative agreement requirements.<sup>5</sup> In general, BLL test results are analyzed at the local and state level for general trends at regular intervals (e.g., quarterly, annually).

Two known temporal patterns commonly occur in CBLs data: seasonal patterns and long-term trends. The seasonal pattern typically shows childhood BLLs peak in summer and late fall.<sup>6,7</sup> The time trend pattern reflects a tendency in most jurisdictions to have a progressively decreasing proportion of child BLLs that exceed 5  $\mu\text{g}/\text{dl}$  over time<sup>8</sup> due to the removal of sources of lead in the environment.<sup>9</sup>

We describe a modified CUSUM/Shewhart alerting algorithm that accounts for seasonal and time trends to assess monthly increases in the proportion of children tested with a BLL  $\geq 5$   $\mu\text{g}/\text{dl}$ . This enhanced surveillance method may allow state and local programs to more efficiently recognize trends that may require public health action.

## Methods

### Algorithm method

Our algorithm applies to a specific jurisdiction, such as a county, and uses children's BLLs, categorized as binary: elevated (a child BLL  $\geq 5$   $\mu\text{g}/\text{dl}$ ) or not elevated (a child BLL  $< 5$   $\mu\text{g}/\text{dl}$ ). Specifically, we aggregated all blood lead test results from child blood lead surveillance data for selected jurisdictions. Test results were aggregated by month and the percent of tests with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$  for each month during the study period (January 2007 to May 2016) were calculated. The method uses 60 consecutive months of historic child blood lead surveillance data to model the observed pattern and to predict the 61st month. The observed value for the 61st month is compared with the predicted value using CUSUM and Shewhart control chart procedures.

If a child had multiple tests (based on a unique identifier), we included the child in the analyses each calendar year in which he/she had a test result. For children with multiple test results in a calendar year, we conducted a sensitivity analysis by employing two approaches: one allowing children's results to be considered annually (i.e., primary analysis) and one allowing children's results to be considered monthly. The first approach retained one test per child per calendar year, keeping the highest venous test. If a child had a capillary or missing test type result, the lowest capillary test result was selected, because of potential upward bias of capillary results.<sup>10</sup> This first approach was considered because it considers new BLL test (i.e., screening test) results  $\geq 5$   $\mu\text{g}/\text{dl}$  and because on average it takes children's BLLs slightly more than 1 year to decline from  $\geq 10$   $\mu\text{g}/\text{dl}$  to  $< 10$   $\mu\text{g}/\text{dl}$ .<sup>11</sup> The second approach retained one test per child per month, keeping the highest venous test. If a child had a capillary or missing test type result, the lowest capillary test result was selected. In both approaches, the monthly denominator was the number of unique children tested for blood lead during

the month. The outcome of interest was the proportion of tests equal to or higher than 5  $\mu\text{g}/\text{dl}$  for each month.

To identify whether the most recent proportion of children with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$  differed substantially from the corresponding model-predicted value, we applied quality control procedures used in industry to child blood lead surveillance data, like other approaches with surgery outcomes.<sup>3,12</sup> The residuals (observed minus expected values) and associated model predicted standard errors from the autoregressive model were used in CUSUM analyses to construct modified CUSUM and Shewhart control charts. We modified the approach by setting the control limit for only the last month (month 61). For month 61, the value is flagged as out of statistical control and as potentially of public health concern if it lay outside the control limits (i.e.,  $\geq 3$  SDs) for either the Shewhart or CUSUM chart control limits. Further, because the observed minus predicted values may not be normally distributed, we recommend empiric adjustment of the CUSUM parameters (shift to be detected [ $\delta$ ], the decision interval [ $h$ ], and the reference interval [ $k$ ]). Adjustment allows the investigator to modify (upward or downward) the number of months expected to be flagged based on historical data and recognized events, if any.

To account for the recognized pattern of declining BLLs over time, we used a regression model that includes a linear time trend, with a join point that allows the pattern to change. Because patterns can change over time, we also restricted analyses to the most recent 61 months, with the join point placed to allow a change in trend over the most recent 24 months of the 61-month period. To account for the seasonal BLL pattern, three knots were used, placed at months 3, 6, and 9 of every year, with a restriction in the regression model so that the model-predicted value at the end of 1 year coincided with the start of the next year. The splines were constrained so that the spline value at months 1 and 13 would coincide (over a 12-month period). We fitted the model using months 1–60 and then calculated the model-predicted value for month 61 (the most recent proportion of children with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$ ). If the observed value for month 61 differed from the predicted value by  $\geq 3$  SDs, this month was flagged as potentially signaling a meaningful change in expected BLL patterns. To adjust for potential correlation over time, we fitted the model using an autoregressive-1 model.<sup>13</sup>

### Method validation

To validate our approach, we used our modified CUSUM/Shewhart alerting algorithm to detect aberrations in de-identified blood lead data transmitted from state and local programs (jurisdictions) to CDC's CBLs system. We compared the jurisdiction-month combinations flagged by the algorithm, with known issues of changes in patterns of child BLLs for two jurisdictions. We applied the algorithm to the two jurisdictions consisting of two "known positive" jurisdictions (i.e., with a known monthly increase in child BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$  during the study period) and 18 randomly selected counties from the CBLs. We randomly selected 18 counties with no known or suspected change in the proportion of children with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$  from states providing continuous childhood BLL surveillance data to CDC during the study period (January 2007 to May 2016). This 113-month study period was used for method validation and to identify historic patterns of the jurisdiction under consideration. Selected counties required for inclusion an average of 10 children with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$  per month during the full study period.

We considered child blood lead surveillance data in successive moving windows of 61 months/window. For example, we analyzed the data assuming surveillance data were available only up to and including January 2012. We used historical data from January 2007 to December 2011 (60 consecutive months) to establish a baseline, fit models, construct control charts and determine if there was a potential problem in month 61 (i.e.,

January 2012). We then moved to the next analytic window (February 2007 to January 2012) and repeated the process. Beginning in January 2013, more months of data were available, so we used 72 consecutive months in our analytic window.

A total of 111 counties from 22 eligible states were available for random selection, and one county per state was selected from eligible states. We extracted CBLs data during October 2016. At the time of selection, 1,547 counties were in the CBLs. Variables included in the analysis were blood lead test result, date of test, child birthdate, age of child at test, unique child ID, and sample type (i.e., capillary/venous/unknown). Additionally, we obtained (via a separate data sharing agreement) blood lead data from two jurisdictions in one state where a known increase in children with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$  had occurred during the study period. The 113-month study period was used to match available data between the randomly selected 18 counties and the two jurisdictions obtained via data-sharing agreement.

Although we did not evaluate this potential modification empirically, we provide the code for a generalized auto-regressive model with conditional heterogeneity (GARCH model) that can account for potentially changing variances over time and incorporate other factors, like number of tests into the variance calculations. We did not pursue this modification for two reasons: first, it would require additional parameters, possibly adding instability to the model; and second, our empiric evaluations (see Results) suggest that the less complicated model works well.

All analyses were conducted using the AUTOREG, CUSUM, and SHEWHART procedures in SAS version 9.3 (SAS Institute Inc., Cary, North Carolina). Starting with SAS default values, we chose parameters for the CUSUM and Shewhart control charts, modifying them based on surveillance data evaluation for an area (Jurisdiction 1) which had a known increase in the proportion of children with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$ . We chose  $\delta = 3$  (shift

to be detected; there is no default, SAS requires input),  $h = 3.0$  (the decision interval; there is no default, SAS requires input), and  $k = 1.0$  (the reference interval; default =  $\text{SD}/2$ ).<sup>14</sup> We chose parameter values so that the “event” in the two “known positive” jurisdictions would be detected. SAS codes are presented in Supplemental Appendix 1; <http://links.lww.com/EE/A78>.

Because the SAS output for the method validation is a combined CUSUM and Shewhart control chart for every window (generating numerous graphics for each jurisdiction investigated), we reconstructed the SAS-developed algorithm using R software (R Foundation for Statistical Computing, Vienna, Austria). The R method reconstructed SAS ARIMA (including the spline calculation), CUSUM, and SHEWHART procedures. R code and quantitative calculations are presented in Supplemental Appendices 2; <http://links.lww.com/EE/A79> and 3; <http://links.lww.com/EE/A80>. Differences in the ARIMA procedures between SAS and R were minimal. Comparing SAS and R output revealed that typically, residuals and standard errors were within 0.01% within the same month (results not shown).

Results were visualized in two ways. The first approach used SAS-generated CUSUM/Shewhart chart output; the second approach employed the R Shiny application to improve visualization and interpretation of results. R Shiny outputted results from each jurisdiction to a single, color-coded, interactive graphic using a web browser. The R Shiny application accepted data in .csv, .xlsx, or .sas7bdat formats. The method constructed the 61st-month windows; performed the reconstructed ARIMA, CUSUM, and SHEWHART SAS procedures on each jurisdiction; and outputted a single visualization, where each point was color-coded to the alert level of when that point corresponded to the final (61st) month of a window. This step reduced the time to interpret results from two visualizations (i.e., observing one CUSUM and one Shewhart chart) per window (96

**Table 1.**  
**Characteristics of 20 selected jurisdictions, January 2007 to May 2016.**

US region	Population of children <6 years (2015) <sup>a,b</sup>	Percent of housing built before 1950 (2015) (%) <sup>b</sup>	Among children <6 years of age						Exceed 3 $\sigma$ threshold	Month of exceedance
			Blood lead tests per month		Blood lead levels $\geq 5$ $\mu\text{g}/\text{dL}$ per month <sup>a</sup>					
			Mean	Median	Mean	Median	%			
West	330,000	3	2,920	2,957	77	37	3	No	NA	
Northeast	2,000	29	413	413	36	35	9	Yes	January 2014	
Southeast	75,000	1	541	530	14	12	3	Yes	October to December 2015	
Midwest	3,000	43	80	77	17	15	22	No	NA	
Midwest	35,000	20	241	217	31	22	13	No	September and October 2015	
Mid-Atlantic	60,000	16	1,589	1,559	66	44	4	No	NA	
Northeast	35,000	39	1,511	1,510	144	112	10	No	NA	
Midwest	60,000	9	1,188	1,156	32	29	3	No	NA	
Midwest	45,000	32	988	999	68	64	7	Yes	December 2013	
South	5,000	3	36	34	17	16	46	No	NA	
Midwest	55,000	25	948	922	39	35	4	No	NA	
Northeast	5,000	26	54	53	18	12	33	No	NA	
Mid-Atlantic	20,000	19	405	419	32	25	8	Yes	December 2015	
Northeast	35,000	24	1,162	1,155	40	35	4	No	NA	
Midwest	10,000	33	288	291	30	25	11	No	NA	
Southwest	70,000	15	571	602	29	27	5	No	NA	
Northeast	45,000	46	1,601	1,633	205	171	13	No	NA	
Midwest	15,000	30	403	396	42	38	11	Yes	June 2012	
Midwest	30,000	19	654	614	21	20	3	Yes	July to October 2014	
Midwest	55,000	23	778	709	90	76	11	Yes	July and August 2015	

Source: U.S. Census Bureau, 2011–2015 American Community Survey 5-Year Estimates, Year Structure Built (Table B25034) [https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS\\_15\\_5YR\\_B25034&prodType=table](https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_15_5YR_B25034&prodType=table) Accessed July 27, 2019.

<sup>a</sup>Source: U.S. Census Bureau, 2011–2015 American Community Survey 5-Year Estimates, Population under 18 Years by Age (Tables B09001003, B09001004, and B09001005) [https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS\\_15\\_5YR\\_B09001&prodType=table](https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_15_5YR_B09001&prodType=table) Accessed July 27, 2019.

<sup>b</sup>Rounded.

NA, not applicable.

visualizations for analyzing 108 months of data), to a single visualization with 108 color-coded data points. This approach combined the monthly proportion of children with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$  with exceedances of the alerting algorithm at 1-, 2-, and 3-SD levels. Users of the R Shiny application have an interface to modify the shift to be detected, decision interval, and reference interval.

**Results**

For the method validation component, in the 20 selected jurisdictions representing 18 states, the average estimated population of children less than 6 years of age was 49,500 (range 2,000–330,000) and the median was 33,400 children <6 years of age. On average, the proportion of pre-1950 housing, a risk factor for lead exposure,<sup>15,16</sup> was 22.8% (range 1%–46%) and the median was 23.4%. All regions of the US were represented except for the northwest (Table 1). Among the 18 randomly selected jurisdictions, we identified alerting signals in six (33%) jurisdictions in 113 (1.3%) months. Among all 20 jurisdictions, 36% of children, on average, had more than one test over the entire study period.

During the study period (January 2007 to May 2016), all 20 jurisdictions had both a downward trend in the monthly proportion of children with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$  and a seasonal (late summer or early fall) increase of children with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$ . Figure 1 demonstrates these two patterns among the four selected counties from the study. We identified alerting signals in the two “known positive” jurisdictions during the same months when documented

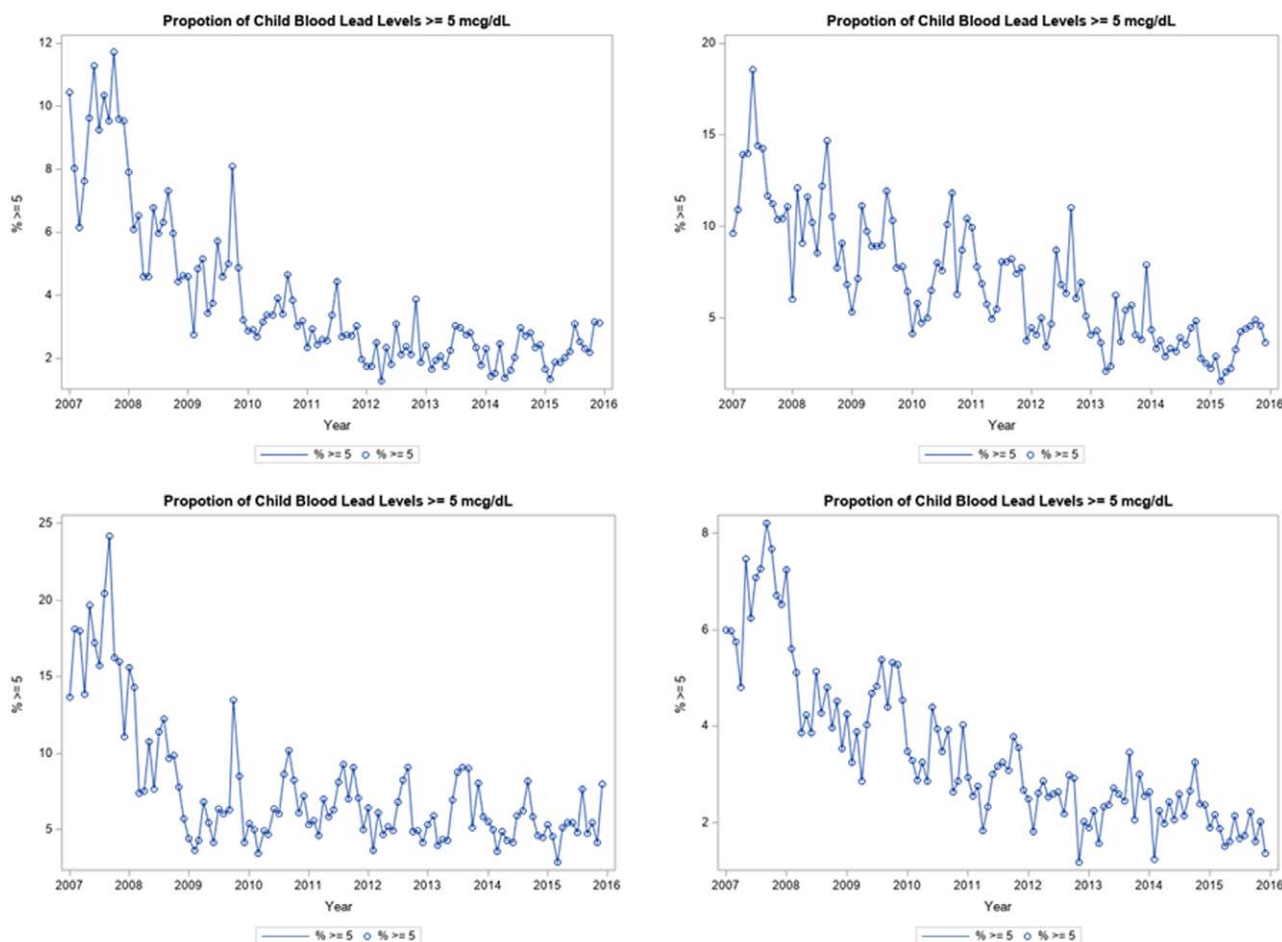
increases in the proportion of children with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$  were known to have occurred. CUSUM and/or Shewhart control chart output for Jurisdiction 1 exceeded the 3-SD threshold from July to October 2014 (Figure 2). In Jurisdiction 2, CUSUM and/or Shewhart control chart output exceeded the 3-SD threshold from July to August 2015 (Figure 3). Improved visualization of results was employed on Jurisdiction 2 (Figure 4).

In sensitivity analyses, 1.4 more children per month, on average, had BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$  when considering highest test per month compared with highest test per year definitions. Using either definition, alerts were raised in the same jurisdictions during the same time periods and via the same control charts.

**Discussion**

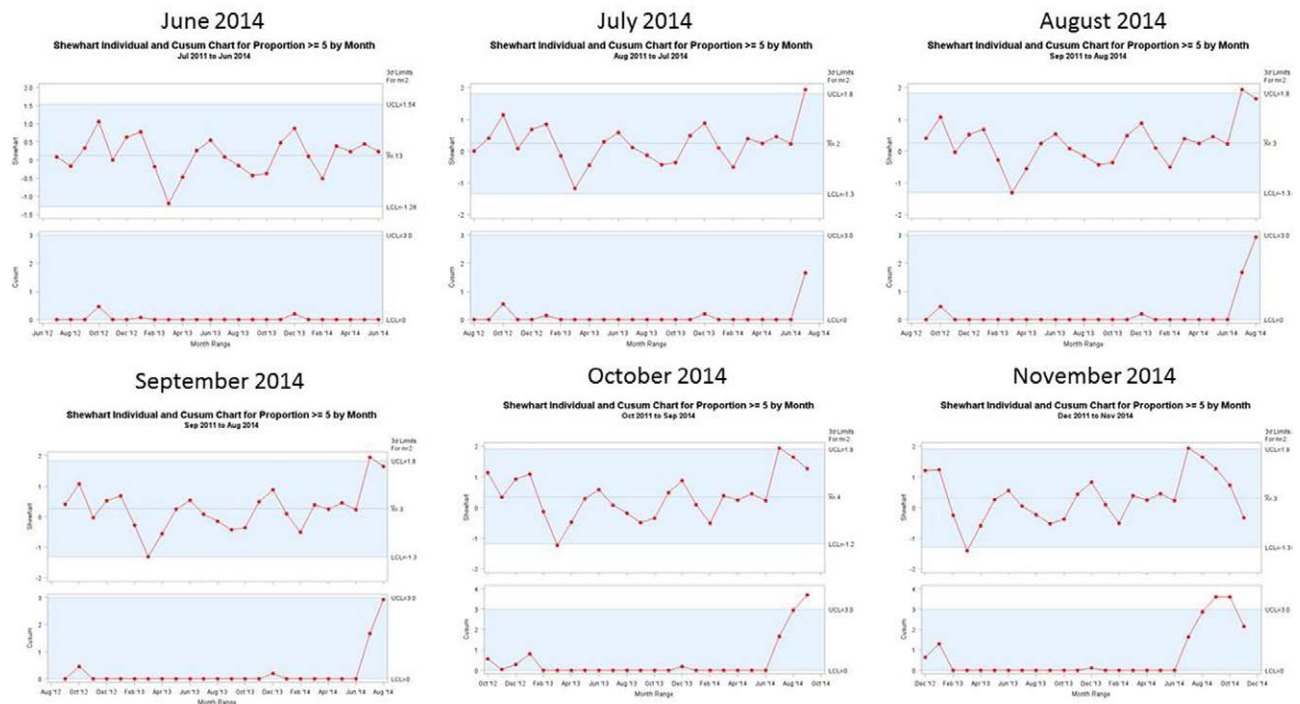
Our modified CUSUM/Shewhart algorithm, never employed, provides a framework for enhanced CBLs, and offers an efficient, rapid secondary prevention approach for identifying changes in the proportion of children with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$ . Alert signals respectively identified time periods in two jurisdictions where a known increase in the proportion of children <6 years of age with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$  occurred. Also, our algorithm accomplished adjustment for seasonality and de-trending over time.

In the two jurisdictions where a known increase of children with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$  occurred, local authorities previously provided follow-up and case management of children based on state and CDC guidance.<sup>17</sup> Among the 18 randomly selected jurisdictions, 13 (72%) did not have alert signals identified. Alerting signals were



**Figure 1.** Proportion of children with blood lead levels  $\geq 5$   $\mu\text{g}/\text{dl}$  by month from CDC CBLs system among four of 18 randomly selected Jurisdictions (Mid-Atlantic, Midwest, South, and Northeast), selected dates: January 2007 to January 2016





**Figure 2.** CUSUM and Shewhart control chart output for Jurisdiction 1, June to November 2014 (3-SD threshold exceedance occurs during July [Shewhart], August [CUSUM], September [CUSUM], and October [CUSUM], 2014).

produced for five (28%) jurisdictions with no known increase of children with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$ . CDC staff examined reporting logs and contacted the respective state child lead poisoning prevention programs to inquire about the changes. Upon further investigation, these alerting signals appear to be related to administrative changes in data management, not in true increases in number of children with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$ . In five of the six (83%) jurisdictions, we identified possible reasons for the alerts: alerting signals at the end of the study period, incomplete reporting, transitioning to new surveillance systems, and submission of previously unsubmitted data by state programs to CDC. In one jurisdiction, an effort to increase blood lead testing in high-risk areas resulted in a 1-month 148% increase in children tested and a 400% increase in children with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$  during a 3-SD alerting signal period. Alerting signals were not impacted by considering the highest monthly test compared to the highest annual test.

To summarize, among the randomly selected 18 jurisdictions, the algorithm identified pattern changes with child BLLs with most of the alerts appearing to be related to data reporting issues and one was a change (increase) in BLL testing. Nonetheless, the alert provides an easy-to-implement and efficient approach to identify deviations in regular patterns of BLLs that require further investigation.

R Shiny (described in the Methods section) can provide childhood lead poisoning prevention program staff a user-friendly means to visualize and interpret complex time series characteristics. A pilot-test for use as a desktop-accessible application is planned. We used this open-source tool to inspect several years of data representing dozens of time windows with a single interactive visualization while retaining alert levels (compared with the original SAS output) within a 0.01% range of the 3-SD threshold.

Most counties (1,436/1,547; 93%) in the CBLs did not meet our stringent inclusion criterion of at least 10 children with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$  consecutively per month from January 2007 to May 2016, so we could not evaluate this method among all programs that submit CBLs data to CDC. However, the 7% of eligible counties represented a sizeable population: 5,517,299 children <6 years of age from 22 states (based on 2015 census estimates).

We expect this algorithm is best suited to geographic areas with larger populations (e.g., county). For jurisdictions with fewer than 10 children with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$  per month, we suggest using alternate methods for investigating patterns of children with elevated blood lead test results. For example, applying the algorithm at a higher aggregated level (i.e., combining jurisdictions with a potential common lead exposure) may allow users to meet our inclusion criterion. Using counts or other measures of central tendency of children with elevated blood lead test results is another possibility, which requires further analyses. Applying the algorithm to very large areas (e.g., statewide) may potentially mask localized changes, but alerts generated on higher-level data could be further investigated manually.

CUSUM control charts are commonly used in industrial and manufacturing process control.<sup>18</sup> However, public health surveillance, syndromic surveillance, and outbreak investigation methods have also applied control chart and temporal adjustment methodology. Similar to our approach, Hutwagner et al.<sup>19</sup> applied a CUSUM algorithm (without model building and seasonal adjustment) to the CDC National Salmonella Surveillance System and, using an expected mean of 5 weeks in the algorithm, were able to detect 29 of 38 of US salmonella outbreaks. Hutwagner et al.<sup>20</sup> later described a seasonally adjusted CUSUM method for bioterrorism syndromic surveillance aberration detection using CDC's Early Aberration Reporting Systems. CUSUM methodology proved useful for real-time monitoring of hospital-acquired invasive aspergillosis infection and for early identification and follow-up of an outbreak.<sup>21</sup> Gomes et al.<sup>22</sup> also used the hospital setting to employ CUSUM, Shewhart, and Exponentially Weighted Moving Average charts to detect nosocomial infection outbreaks. The authors concluded that the three charts used in conjunction were useful for detecting nosocomial infection outbreaks and if results are communicated rapidly to hospital staff, may lead to prevention of outbreaks.

This study is subject to limitations. The complex methodology to adjust for the overall downward trend in childhood BLLs, seasonality of childhood BLLs, and autocorrelation requires several years of continuously collected data. Local



**Figure 3.** CUSUM and Shewhart control chart output for Jurisdiction 2, June to September 2015 (3-SD threshold exceedance occurs during July [Shewhart] and August [CUSUM and Shewhart], 2015).

lead poisoning prevention programs may not have appropriate blood lead surveillance data to apply the methodology or the statistical support to modify the methods to meet local conditions. However, we have developed an R Shiny application that incorporates the methodology and is available for free. Additionally, we were restricted to assessing the potential changes in the proportion of children with BLLs  $\geq 5$   $\mu\text{g}/\text{dl}$  as the outcome of interest. Our analysis might be enhanced by assessing certain measures of central tendency (e.g., changes in mean/median BLLs). Because of differences in reporting limits of clinical laboratories, including users of point-of-care blood lead analyzers, we did not assign values to BLL results  $< 5$   $\mu\text{g}/\text{dl}$  because it does not accurately reflect real-world conditions. We would have liked to validate our algorithm method in more than two jurisdictions with documented community lead exposure during a specific time period. After we developed the algorithm, we contacted state partners to apply it on additional child blood lead surveillance data from specific jurisdictions. However, jurisdictional lead exposure was not possible to be confirmed during a specific time period by our partners, thus not allowing for our algorithm to be further validated. Finally, the algorithm has the potential for alerting notifications of nonpublic health importance because of its high sensitivity. We identified alerts resulting from data reporting issues rather than actual changes in children’s lead exposure, as measured by BLLs. However, the

algorithm is designed to alert public health officials to potential changes requiring further investigation.

**Summary**

CDC and partners support primary prevention—the control or removal of sources of lead before children are exposed. However, this new secondary prevention approach provides a framework for enhanced surveillance of childhood blood lead data and an opportunity for public health officials to rapidly investigate important trends in exposure patterns. Further evaluation of the algorithm in real-world conditions by local and state health departments can identify and evaluate alert settings of public health significance. We provide the practitioner with SAS and R code in the appendices.

**ACKNOWLEDGMENTS**

We thank Martha Stanbury, Division of Environmental Health, Michigan Department of Health and Human Services for her review and comments on the manuscript. We also thank Nadonna Jones, Office of Communication, National Center for Environmental Health/Agency for Toxic Substances and Disease Registry for her copy editing. Finally, we thank Ted Larson, Environmental Health Surveillance Branch, Agency for



**Figure 4.** CUSUM and Shewhart control chart output for Jurisdiction 2 using R-Shiny Application, January 2007 to May 2016 (3-SD threshold exceedance occurs during July and August 2015)

Toxic Substances and Disease Registry, for his assistance with SAS programming.

## References

- Hanna-Attisha M, LaChance J, Sadler RC, Champney Schnepf A. Elevated blood lead levels in children associated with the Flint drinking water crisis: a spatial analysis of risk and public health response. *Am J Public Health*. 2016;106:283–290.
- Ruckart PZ, Ettinger AS, Hanna-Attisha M, Jones N, Davis SI, Breyse PN. The Flint water crisis: a coordinated public health emergency response and recovery initiative. *J Public Health Manag Pract*. 2019;25(suppl 1):S84–S90.
- Grigg OA, Farewell VT, Spiegelhalter DJ. Use of risk-adjusted CUSUM and RSPRT charts for monitoring in medical contexts. *Stat Methods Med Res*. 2003;12:147–170.
- Centers for Disease Control and Prevention. Prevent Childhood Lead Poisoning Prevention. Infographic, 2013. Available at: <http://www.cdc.gov/nceh/lead/publications/lead-infographic-final-full.pdf>.
- Pertowski C. Lead poisoning. In: *From Data to Action: CDC's Public Health Surveillance for Women, Infants, and Children*. Atlanta, GA: Centers for Disease Control and Prevention, US Department of Health and Human Services; 1994:311–319.
- Yiin LM, Rhoads GG, Liyo PJ. Seasonal influences on childhood lead exposure. *Environ Health Perspect*. 2000;108:177–182.
- Laidlaw MA, Mielke HW, Filippelli GM, Johnson DL, Gonzales CR. Seasonality and children's blood lead levels: developing a predictive model using climatic variables and blood lead data from Indianapolis, Indiana, Syracuse, New York, and New Orleans, Louisiana (USA). *Environ Health Perspect*. 2005;113:793–800.
- Raymond J, Brown MJ. Childhood blood lead levels in children aged <5 years - United States, 2009–2014. *MMWR Surveill Summ*. 2017;66:1–10.
- Dignam T, Kaufmann RB, LeSturgeon L, Brown MJ. Control of lead sources in the United States, 1970–2017: public health progress and current challenges to eliminating lead exposure. *J Public Health Manag Pract*. 2019;25(suppl 1):S13–S22.
- Caldwell KL, Cheng PY, Jarrett JM, et al. Measurement challenges at low blood lead levels. *Pediatrics*. 2017;140:e2017027.
- Dignam TA, Lojo J, Meyer PA, Norman E, Sayre A, Flanders WD. Reduction of elevated blood lead levels in children in North Carolina and Vermont, 1996–1999. *Environ Health Perspect*. 2008;116:981–985.
- Lovegrove J, Sherlaw-Johnson C, Valencia O, Treasure T, Gallivan S. Monitoring the performance of cardiac surgeons. *J of the Oper Res Soc*. 1999;50:684–689.
- Rothman K, Greenland S, Lash TL. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008:572–573.
- SAS Institute Inc. SAS/QC® 15.1 User's Guide The CUSUM Procedure. Copyright © 2018, SAS Institute Inc., Cary, NC, USA. All Rights Reserved. Produced in the United States of America. 2018. Available at: <https://support.sas.com/documentation/onlinedoc/qc/151/cusum.pdf>. Accessed 24 July 2019.
- Oyana TJ, Margai FM. Geographic analysis of health risks of pediatric lead exposure: a golden opportunity to promote healthy neighborhoods. *Arch Environ Occup Health*. 2007;62:93–104.
- Gaitens JM, Dixon SL, Jacobs DE, et al. Exposure of U.S. children to residential dust lead, 1999–2004: I. housing and demographic factors. *Environ Health Perspect*. 2009;117:461–467.
- Centers for Disease Control and Prevention. Managing Elevated Blood Lead Levels Among Young Children: Recommendations from the Advisory Committee on Childhood Lead Poisoning Prevention. Atlanta: CDC; 2002. Available at: <https://www.cdc.gov/nceh/lead/casemanagement/managingEBLLs.pdf>. Accessed 24 July 2019.
- Montgomery D. *Introduction to Statistical Quality Control*. Hoboken, N.J: Wiley; 2001.

19. Hutwagner LC, Maloney EK, Bean NH, Slutsker L, Martin SM. Using laboratory-based surveillance data for prevention: an algorithm for detecting Salmonella outbreaks. *Emerg Infect Dis.* 1997;3:395–400.
20. Hutwagner L, Thompson W, Seaman GM, Treadwell T. The bioterrorism preparedness and response Early Aberration Reporting System (EARS). *J Urban Health.* 2003;80(2 suppl 1):i89–i96.
21. Menotti J, Porcher R, Ribaud P, et al. Monitoring of nosocomial invasive aspergillosis and early evidence of an outbreak using cumulative sum tests (CUSUM). *Clin Microbiol Infect.* 2010;16:1368–1374.
22. Gomes IC, Mingoti SA, Oliveira CD. A novel experience in the use of control charts for the detection of nosocomial infection outbreaks. *Clinics (Sao Paulo).* 2011;66:1681–1689.