A Public Health Performance Evaluation Primer

From the University of Minnesota Public Health Preparedness and Emergency Response Research Center: University of Minnesota: Simulations, Exercises, and Effective Education

U-SEEE: UNIVERSITY OF MINNESOTA: SIMULATIONS, EXERCISES, AND EFFECTIVE EDUCATION

PERRC Preparedness & Emergency Response Research Center

This work is sponsored by University of Minnesota: Simulations, Exercises and Effective Education Preparedness and Emergency Response Research Center (U-SEEE PERRC). U-SEEE Principal Investigator: Debra K. Olson. This work is part of "The Retrospective Cohort Study of Responders Training and System Performance" Lead Investigator: Craig Hedberg, PhD. Funding made possible, in part by, grant #5P01TP000301-04 from the Centers for Disease Control and Prevention (CDC). The content is the sole responsibility of the authors and does not necessarily represent the official views of the CDC.

What is performance evaluation?

Performance evaluation is the use of quantitative data generated by public health programs to describe how often and how well public health activities are carried out. Performance evaluation is similar to a self-audit of public health program activities, and can be used to improve performance within health departments.

Why should public health programs conduct performance evaluation?

The first reason public health programs should conduct performance evaluation is to improve the quality of programs. If an agency is unaware of the level of performance of a program, it will not be able to measure improvement over time. Performance measurement is also necessary for public health accreditation, public accountability, and giving public health staff a sense of ownership of the quality of their work. Providing performance measurement results to funding authorities may provide justification for continued and increased investment in public health programs.

How is a performance evaluation started?

The first step towards beginning a continuous performance evaluation process is forming a performance evaluation team. The team may be composed of existing staff who dedicate a few hours per week to measuring program performance. The team should meet regularly with management to agree upon performance measurement methods and identify processes to measure.

How is performance evaluation conducted?

There is no right or wrong way to measure performance. We suggest that there are 6 basic steps in continuous performance measurement:

1. Define the program goals and processes
2. Define the process performance metrics
3. Measure the process
4. Evaluate the performance of the process with Run Charts
5. Identify strategies for improvement, if possible
6. Improve processes with the Plan-Do-Check-Act cycle

Performance Evaluation Process Details

1. Define the process and program goals

The first task of the performance measurement team is to identify the goals of the public health program, and then identify the processes that contribute to the performance of the program should be described. This usually involves interacting with all program staff to develop process maps (Figure 1) to describe each step in the processes conducted by the program, as well as the internal and external stakeholders that contribute to each process. For example, the process of sending a stool sample from a local health department to the state health department involves

several recordkeeping steps and three stakeholders: the local health department, the courier, and the state health department. See the example at the end of this document.

2. Define the performance metrics

After the program goals and processes that contribute to the program goals have been described, the next step is to determine how the processes should be performing. Public health programs may choose to evaluate processes based on how frequently processes are completed, home completely processes are completed, or how quickly processes are completed, depending on the goals of the public health program. For example, the Minnesota Department of Health is responsible for conducting food exposure history interviews with cases of Salmonella reported to the health department. In one evaluation, MDH wanted to measure how long it took to complete the interviews after receiving a report of a case, so the department selected to measure the number of days that elapsed between the receipt of a laboratory isolate of Salmonella and the interview that occurred with the ill person.

3. Measure the process

After the processes have been described and the performance metrics selected, data must be abstracted from public health programs in order to evaluate performance over time. This is usually the most time-intensive part of the evaluation process. Evaluations are often conducted using retrospective data, and the evaluation team must identify where key performance data is recorded and abstract it. Data is often found in paper forms that record key dates and or information that is needed in order to meet public health program goals (such as a food history obtained in an interview with a person who reports a foodborne illness). It is possible that data may not exist and that data collection systems must be setup to measure data prospectively. In this case, a performance evaluation should be completed after a considerable amount of data has been collected.

The key to continuous evaluation is the ability to measure process performance over time. Programs conducting an evaluation should record key measures of timeliness or completeness for every iteration of the process being measured, depending on the goals of the program. This information can be condensed into monthly averages for process performance, or the entirety of process performance data can be evaluated on a Run Chart, described in the next section.

4. Evaluate the performance of the process using Run Charts

Once performance data has been collected, we recommend that the data be evaluated using a chart known as a Run Chart. Run Charts plot process performance over time, with the performance data shown on the Y axis, and process iterations or weekly/monthly averages plotted on the X axis. Once the data is plotted, the median of the data is calculated and overlaid on the chart. See Figure 2 for an example of a Run Chart.

Run Charts use statistical rules to determine time periods in process performance where performance is unexpectedly bad or good. The idea with using run charts to measure performance over time is that once periods of good and bad performance have been identified, brainstorming sessions can be used to identify reasons performance behaved the way it did over these time periods.

Once a run chart has been completed, we recommend that the performance evaluation group visually analyze the chart to identify areas of good and bad performance. There are 4 rules the performance data should follow if there is no indication of good or bad performance (in other words, the process is performing in a "stable" manner). The four rules are:

1. The performance data line should cross the median line a certain number of times, depending on the number of data points on the chart. (Appendix A)
2. There should be no "runs" of 8 consecutive data points above or below the median line.
3. There should be no "trend" of 6 consecutive data points all increasing or decreasing in a row, regardless of if the "trend" crosses the median line.
4. There should not be a "zig-zag" pattern in which the performance line alternately crosses above and below the median line 12 times in a row.

Run Chart theory says that if rules 1, 3, or 4 are broken by the performance data, then there is a "special cause", or identifiable cause that is causing the process to act in an unstable manner. If one goal of the program is to provide consistent and predictable performance, a brainstorming session may be initiated to determine potential causes of unstable performance.

If rule 2 is broken and long runs above or below the median lines are detected, those are indications of time periods in which performance was either good or bad. Depending on the performance metric being evaluated, a run above the median line could be interpreted as either good or bad performance. For example, if completeness of foodborne illness complaints is being measured, then a run above the median could be interpreted to mean that the completeness of case reports was unusually high during that time period, and that a brainstorming session could be initiated to determine why performance was so good at that time. On the other hand, if the time to complete a foodborne illness complaint was being measured, a long run above the median could indicate that it took an unexpectedly long time to complete the complaint reports during the time period, and a brainstorming session could be initiated to determine why it took so long to complete the complaint reports.

5. Identify strategies for improvement, if possible

If periods of good or bad performance are identified on the Run Chart, the performance evaluation group should begin a brainstorming session, or "root cause analysis" to determine factors contributing to time periods of good or bad performance. If the conclusions of the root cause analysis points to a correctable factor causing bad performance, the program may consider taking steps to change the process in order to reduce the probability of bad performance occurring in the future. If a root cause analysis is conducted to investigate a period of good performance, the program may consider changing the process to increase the likelihood of good performance in the future.

6. Improve processes with the Plan-Do-Check-Act cycle

The Plan-Do-Check-Act (PDCA) cycle is a process used to measure changes in performance over time using Run Charts. Once Run Charts and root cause analysis have been used to evaluate past performance, the PDCA cycle can be used to improve performance using 4 steps:

1. Plan what will be done to the process to improve performance in the future
2. Do an intervention, or change the process
3. Check how the process performance is changing over time on a Run Chart
4. Act to solidify the change in process into routine operations if an improvement in performance is sustained over time on the Run Chart

It is possible that a root cause analysis will identify factors contributing to bad performance that are outside of the influence of the public health program. In this case, no changes to the process can be made that can improve the process performance. The PDCA cycle should only be used when a root cause analysis identifies changeable factors as the cause of exceptionally good or bad performance.

How many resources do we need for performance measurement?

Performance measurement projects can vary in size and scope. Large, comprehensive evaluations can take weeks and demand dozens of hours of data collection, while smaller and simpler projects can be completed with nothing more than Microsoft Excel and a few hours of work each week. Continuous performance measurement can be time consuming and should be based on critical problems and processes. We suggest that public health programs begin a performance evaluation effort with a small project, for which data is easily available. It is important that the time devoted to performance measurement does not actually detract from the actual performance of public health programs.

What are some additional resources on Performance Evaluation and Run Charts?

- Bialek R, Duffy GL, Moran JW. The Public Health Quality Improvement Handbook. Milwaukee: ASQ Quality Press; 2009. 451 p.
- Duffy GL, Moran JW, Riley WJ. Quality Function Deployment and Lean-Six Sigma Applications in Public Health. Milwaukee: ASQ Quality Press; 2010. 196 p.

Case Study and Example Performance Measurement Figures

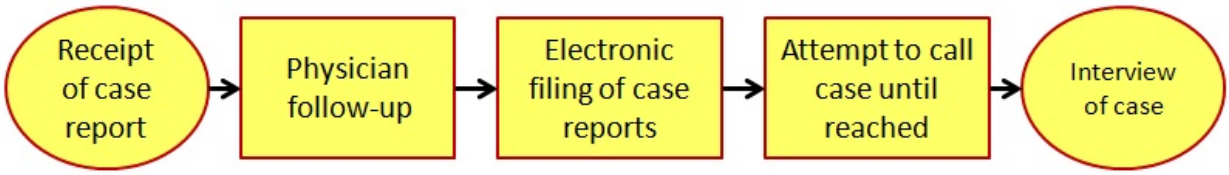
In 2012, The disease X epidemiology unit in state "A" began a performance evaluation to determine the length of time it took the health department to interview cases of disease X about past exposures after the health department had received a case report of a diagnosis from physicians within the state. The following description of the evaluation process used may be a useful illustration of how performance evaluation can be applied in other public health programs.

The leadership of state "A" recruited a student at the local university to assist in the performance evaluation. The student worked with the epidemiologists in the unit who were responsible for completing interviews of persons diagnosed with disease X and reported to the health department.

1. Define the program goals and processes

The evaluation team first made note of one goal of the disease X surveillance program, to interview ill persons as soon as possible in order to obtain illness and previous exposure histories. This goal contributed to the health department's mission of detecting and investigating outbreaks of disease X as quickly as possible in order to prevent illnesses. The evaluation team worked together to describe the process that occurs within the department between the receipt of a case report of disease X and the interview of the case. The team chose to describe the process steps in graphic form in the process map shown in Figure 1 below.

Figure 1. A process map of disease X follow-up at health department A.



2. Define the process performance metrics

The disease X surveillance unit was concerned with measuring the timeliness of the process and therefore chose to evaluate the number of days that elapsed between the receipt of a case report and the interview of the case, as shown on the process map in Figure 1.

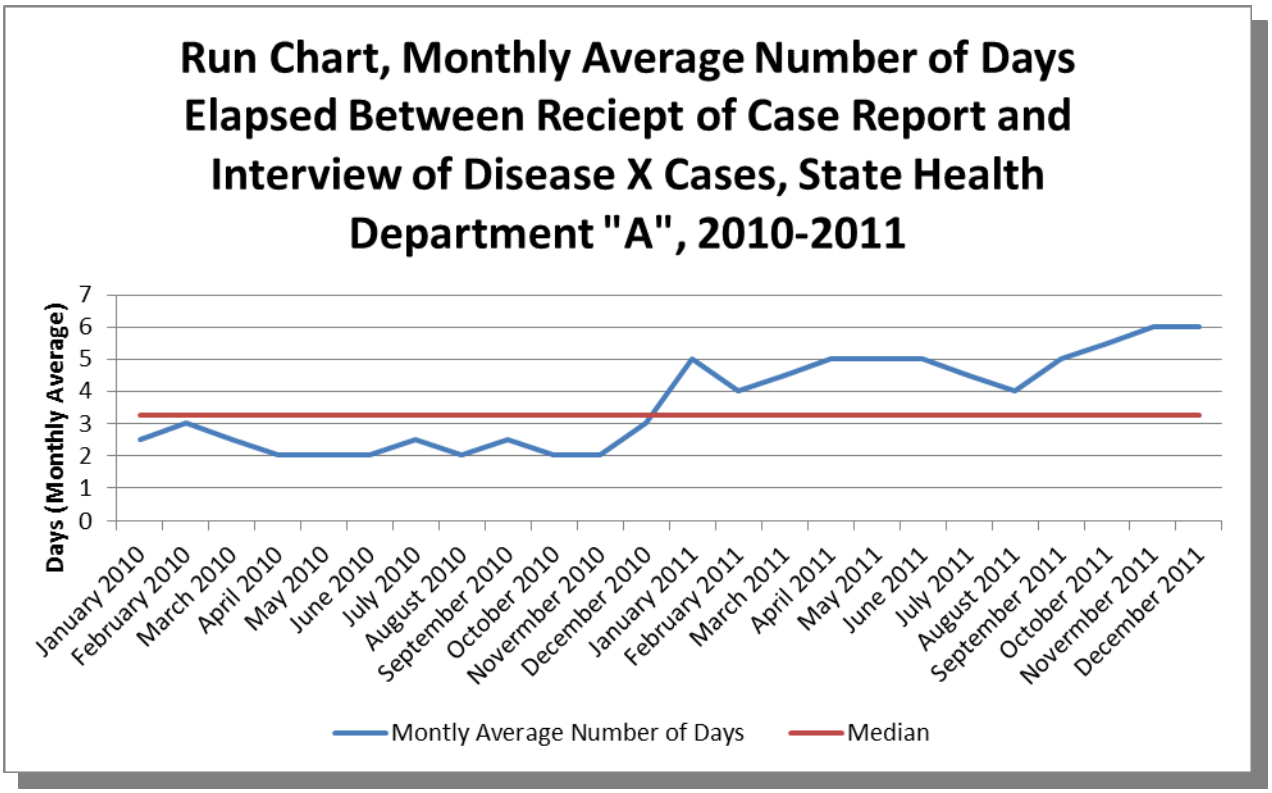
3. Measure the process

The evaluation team abstracted the dates of case report receipt and the date of interview completion from paper copies of case reports and interviews filed within the surveillance unit in the previous 24 months. The dates were entered into an electronic spreadsheet and the number of work days that elapsed between the two dates was calculated.

4. Evaluate the performance of the process with Run Charts

Graphing software was used to calculate the average number of days per month it took to complete an interview after the receipt of a case report. These values (24 total, one average for each month) were plotted on a Run Chart over time. The median of the 24 average values was calculated on overlaid on the Run Chart.

The Run Chart below was analyzed to determine whether the process was stable over the 24 months prior and whether or not there were periods of good or bad performance that could be investigated.



The evaluation team visually analyzed the Run Chart for stability and periods of good and bad performance. The team concluded that the process was not stable because of violations to rules 1 and 2. There were not enough runs observed on the chart, and two long runs were identified; one below and one above the median.

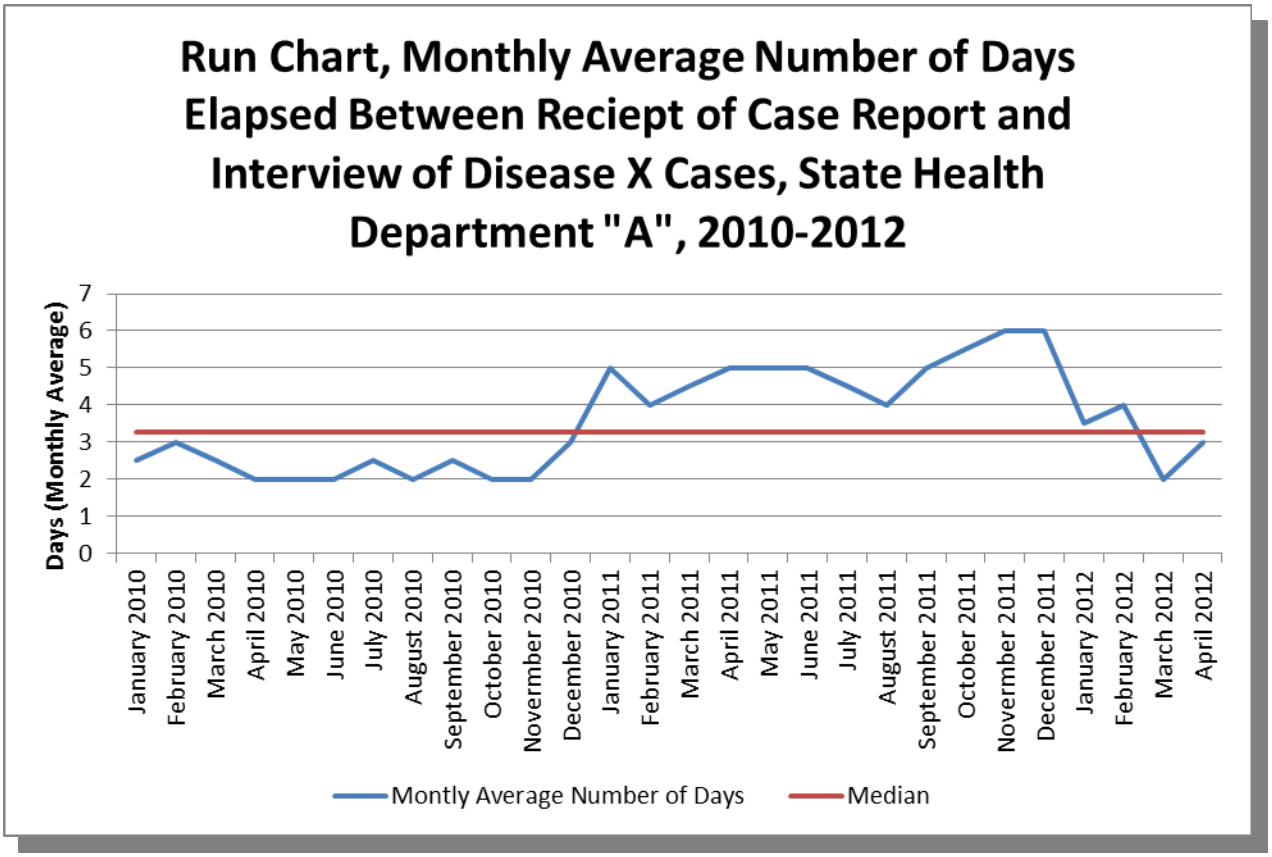
5. Identify strategies for improvement, if possible

A "root cause analysis" brainstorming session was initiated to investigate why the timeliness of the interviews had apparently increased over time. Several causes that could explain the decrease in performance were noted in the root cause analysis, including the hiring of a new interviewer, the implementation of an unfamiliar electronic reporting system, an increase in the total number of reported cases of disease X in the year 2011, and the acquisition of additional job duties by the interviewer beginning in 2011. Consensus was reached by the program staff that the most likely cause of the decrease in performance was due to the interviewer spending time focusing on the newly assigned job duties, combined with an increase in the total number of disease X cases reported during the year 2011.

6. Improve processes with the Plan-Do-Check-Act cycle

In response to the Run Chart Analysis and the root cause analysis brainstorming session, the evaluation team recommended to the program manager that the new job duties acquired by the disease X case interviewer in 2011 be reassigned to another employee within the unit, so that the interviewer could interview cases in a timelier manner. The job duties were reassigned and performance was tracked over time for the next 4 months to check the results of the

intervention on the process. Four months later, the Run Chart began to show signs of improvement in the performance of the process, as shown below.



For more information on how to complete performance evaluations in public health program quality improvement, see the resources listed on page 5 of this brochure.

Appendix A – Number of Runs Expected on Run Charts

Number of Data Points	Least Number of Runs Expected	Most Number of Runs Expected
10	3	8
11	3	9
12	3	10
13	4	10
14	4	11
15	4	12
16	5	12
17	5	13
18	6	13
19	6	14
20	6	15
21	7	15
22	7	16
23	8	16
24	8	17
25	9	17
26	9	18
27	9	19
28	10	19
29	10	20
30	11	2
31	11	21
32	11	22
33	11	22
34	12	23
35	13	23
36	13	24