Quality Improvement Primers

Measurement for Quality Improvement

Ontario

Health Quality Ontario

ACKNOWLEDGEMENTS

This workbook is the result of the efforts of the Health Quality Ontario (HQO) For additional information about other resources, contact: Health Quality Ontario www.hqontario.ca

Individuals may reproduce these materials for their use provided that proper attribution is given to the appropriate source. The recommended citation for this resource guide is: Health Quality Ontario (April 2013).

HQO is funded by the Ontario Ministry of Health and Long-Term Care (MOHLTC).

Table of Contents

OVERVIEW OF MEASUREMENT

Measurement in quality improvement allows a Quality Improvement (QI) team to demonstrate current performance (or baseline), set goals for future performance, and monitor the effects of changes as they are made. Successful measurement is a cornerstone of successful improvement. How do you know if the changes you are making are leading to improvement? Simple – you measure. Measurement does not have to be difficult or time-consuming. The key is to pick the right measures so that the quality improvement team can see results quickly and are able to adapt their interventions accordingly, putting less strain on resources and more focus on outcomes.

Of the many types of measurement, two are typically utilized in health care: measurement for research and measurement for improvement. Traditionally, health care providers are trained to look at research; however, measurement for research is very different from measurement for improvement. The differences are outlined in the table below.

	Measurement for Research	Measurement for Learning and Process Improvement
Purpose	To discover new knowledge	To bring new knowledge into daily practice
Tests	One large "blind" test	Many sequential, observable tests
Biases	Control for as many biases as possible	Stabilize the biases from test to test
Data	Gather as much data as possible, "just in case"	Gather "just enough" data to learn and complete another cycle
Duration	Can take long periods of time to obtain results	"Small tests of significant changes" approach accelerates the rate of improvement

EFFECTIVE MEASUREMENT IN QUALITY IMPROVEMENT

In order to recognize when we have achieved our goals, it is important to define what our 'better' state looks like, and measure to know if the changes we make result in the improvements we seek. The best approach is through the measurement of items called indicators, performance measures or measures.

Begin brainstorming change ideas and associated measurements by asking the following questions:

- What are the current outcomes?
- What outcomes are desired?
- What are the processes and activities that have an impact on the outcomes?
- How are these processes performing?
- Are the processes stable and reliable?
- What will the impact of one process change be on the outcomes?
- What will the impact be on other parts of the system if one process is changed?

All of these questions require a comprehensive understanding of the system or how the various processes work together to achieve outcomes. In quality improvement, this is called a "family of measures," which provide a view of the system from the outcomes, to the processes, to the unintended impacts.

STRATEGIES FOR SUCCESSFUL MEASUREMENT

The strategies below will help the improvement team bring these measurement guidelines to the improvement projects underway in your organization.

Develop a measurement plan

Convene the improvement team and agree on the following:

1. Name of measure
2. Type of measure (outcome, process or balancing, see below)
3. Why the measure is needed for the project
4. Operational definition
5. Data collection and sampling method
6. How will data be displayed?
7. Is baseline data available?
8. Is there a goal or target?
9. Source

What will be measured, how often will it be measured, who will be responsible for measurement, and how will the measurements be shared with the team, leadership and the organization are important questions to answer. Review this plan with the team regularly to ensure that it is working and that there is clarity about what the team is trying to achieve by measuring their progress. Also, make sure that the data collected and analyzed is shared with all stakeholders. It is difficult to create momentum among staff without providing them with relevant and timely information.

Examples of outcome measures:

- *For diabetes: Average hemoglobin A1c level for population of patients with diabetes.*
- *For access: Number of days to third next available appointment.*
- *For critical care: Intensive Care Unit (ICU) percent unadjusted mortality.*
- *For medication systems: Adverse drug events per 1,000 doses.*

Examples of process measures:

- *For diabetes: Percentage of patients whose hemoglobin A1c level was measured each quarter in the past year*
- *For access: Average daily clinician hours available for appointments.*
- *For acute care: Percentage of patients for whom a LACE score was calculated. (LACE stands for: Length of Stay, Acuity of admission, Comorbidities [as measured by a Charlson Score] and number of previous ED visits)*

Use multiple measures.

Consider each category of measures listed below. Collect and plot the data on these measures to create a "family of measures." This will ensure that you have an accurate picture of the effects of the changes your quality improvement team has tested.

Outcome measures are the "voice of the patient or customer" and capture system performance. They answer the question: "What are the end results of our QI work."[1]

Process measures are the "voice of the workings of the system." In other words, process measurements are those that capture the changes your quality improvement efforts make to the inputs or steps that contribute to system outcomes. When working with process measures, it is important to focus on the processes that directly contribute to the outcome that is desired.[2]

A good example of linking the process measure to the outcome is measuring the percentage of time staff comply with a best practice recommendation that will prevent a negative patient outcome (for example, bundle compliance with the 'bundle' of best practices to prevent ventilator associated pneumonia).

Balancing measures determine whether changes designed to improve one part of the system are causing new problems in other parts of the system. For example, does this new QI change improve staff satisfaction but decrease client satisfaction?

PDSA Measures are those that are collected with each test of change (PDSA) that is carried out. These measures provide information about the effect of each change attempt.

PLAN-DO-STUDY-ACT (PDSA)

Step 1	PLAN (who, what, where, when, and why)
	• State the purpose of the PDSA—are you developing a change idea, testing a change, or implementing a change? • What is your change idea? • What indicator(s) of success will you measure? • How will data on these indicators be collected? • Who or what are the subjects of the test? • How many subjects will be included in the test and over what time period? • What are your predictions as to what will happen?
Step 2	**DO**
	• Conduct the test. • Document the results, measurements, challenges and unintended consequences.
Step 3	**STUDY**
	• Analyze the data and study the results. • Compare the data to your predictions. • Summarize and reflect on what was learned.
Step 4	**ACT**
	• Refine the change idea based on lessons learned from the test. • Prepare a plan for the next test. Dependent on results the idea should be adopted, adapted or abandoned.

LEADING AND LAGGING INDICATORS:

“Leading” and lagging” indicators offer more ways of gaining knowledge about how a system is performing and where to focus your quality improvement efforts.

A leading indicator provides information about how a process is performing after changes have been implemented. A lagging indicator is one that provides information about how the system is performing after changes have been made. In quality improvement, process measures are usually referred to as “leading,” while outcome measures are referred to as “lagging”. If the performance of a process measure begins to drop, it is likely that the performance of an outcome measure will also decline.[3]

For example, assessment of residents in long-term care for falls (process) will typically serve to prevent falls (outcome), as actions are taken in response to what is learned in the assessment. If the quality improvement team focused on the level of completion of these assessments over time and discovered that the rate of completion declined, it would be a fairly good predictor that the level of falls would increase. Constantly measuring and monitoring process measures can help the team to understand what may be causing a decline in outcomes, and can also help the team avoid negative outcomes before they happen.

Choose appropriate measures to understand your system.

Raw data is hard to compare. Statistics are used to organize and summarize the information that is collected. The basic summarizing statistics that are likely used in your quality improvement efforts are:

Counts: A count of how many items or observations
Example: the number of people responding to a survey

Sums: Adding up the number of items or observations
Example: 20 out of 100 people surveyed feel that communication with their healthcare provider is inadequate

Ratio: A fraction that describes two groups relative to one another.
Example: the ratio of females to males in the study

Rate: A ratio that describes one quantity in relation to a certain unit.
Example: the rate of infection expressed per 1000 patients

Rates, ratios and percentages help you standardize your data so that it is expressed in a meaningful way that can be readily compared with other data. Ratios and rates may be expressed as percentages. How you choose to present your data will depend on the nature of your data and how you plan to use it.

Ratios (percentages) may be used to adjust for the impact of natural changes in your system, such as volume. The numerator is the key measure (e.g., costs, patients waiting) and the denominator is the unit of production or volume (e.g., total costs, total patients waiting). For example, if the number of patients waiting for more than one hour increased dramatically, you might draw one conclusion. If you knew that overall volume had also increased (which would show up in the ratio), you would mostly likely draw another, more accurate, conclusion.

Sometimes the denominator may be so large that the change looks imperceptible. For example, one year the rate of Ventilator Acquired Pneumonia (VAP) in a hospital was 0.13% and the following year it was 0.26%. That rate is less than one percent of the patients treated. Therefore, it rose slightly, but not by much. To look at your system from a more detailed perspective, use whole numbers, which provide more information about what is happening on a daily basis, without the “watering-down” effect of a ratio. Real numbers speak more directly to the experience of your customers. How many customers are experiencing "x"? In the case of the VAP example - in the first year, two people acquired VAP and the following year four people did. Since VAPis almost always preventable, an additional two people experiencing a secondary, needless hospital induced illness, while suffering from a critical illness, is serious. Although it is useful to understand an increase in volume (as in the above example of patients waiting), it is crucial to understand the experience of each patient when conducting improvement work.

Integrate measures into your daily routine.

To see if changes are leading to improvement, use the "get just enough data" approach. Where measurements are gathered by auditing patient/client/resident charts, for example, decide how many charts will provide enough data for a fairly accurate depiction of your system and consistently use that number to measure over time. Graph and display your measures often enough to give your team feedback in a timely manner, both to keep up momentum and to learn of changes that are having adverse effects. Monthly graphs are often best suited for larger outcome measures whereas weekly graphs may be preferable for smaller, more variable process measures. In your quality improvement team, ensure that there is enough time allotted for staff to review the results and plan the next steps.

Try to build data collection into the daily routine instead of making it a separate project. This not only ensures that data is timely but also reduces stress by making measurement something that is simple to do. Create data collection forms that include only the information you need and that are easy to fill out. When integrating measurement into a staff member's role, be sure to build in a contingency plan for ongoing collection should that person be unavailable.

Plot data over time.

Although one of the more common ways to collect and display data is the pre/post method (i.e., collect data before and after a change to the system or process), displaying data in bar charts is of limited value to improvement efforts because it does not answer the question "What are the effects of making this change?" Summary statistics can hide information about outliers and patterns. In improvement efforts, changes are not fixed but continuously adapted over time.

The best way to collect and display data is to use run charts and statistical control charts –graphical records of a measure plotted over time (often months). Charts annotated with changes and events provide even more information and can help you more accurately make connections between interventions/events and outcomes.

When measuring effects or incidents that are extreme but episodic (for example, the outbreak of an illness), tracking the time between episodes will give you more useful information.

Develop visual displays of measures.

Visual displays are communication tools, motivators, reality checks, and validations of work already done. They don't need to be perfect, just useful. And don't wait until information systems are ready - start with simple data collection methods such as paper and a pencil. Having clearly visible data will point the improvement team in the right direction.

Build in ongoing measurement and communication

Establishing an ongoing measurement system and a standardized way of communicating results reinforces the idea that change is important to the organization. By creating and sharing the collected data, the quality improvement team is likely to gain the support of the organization as a whole.

SAMPLE MEASURES

When collecting measures for a quality improvement project, teams often ask how much they should collect. Should every patient chart be examined and recorded? Should every element that touches the outcome be measured? The very simple answer is to measure enough to create knowledge and understanding about the system. Each system is different, so whether examining five charts is enough or it takes ten, it is important not to overwhelm the improvement team by creating so much work that it cannot get the measurements done. It is, however, crucial to ask whether what is being measured gives enough information to understand how the system is performing and how it will react to planned changes. Once "how much" has been determined, it is necessary to determine "how often," still keeping in mind the availability of the information and of the team to carry out the measurement.

Using Sampling: An Example

Here is how one team used sampling to measure the time for transfer from Emergency Department (ED) to inpatient bed. Rapid movement from the ED after the decision to admit a patient is critical to the entire system for emergent patient care.

Sampling Approach

The measurement consisted of six weekly data collections of 25 patients each. The patients were sampled in several ways:

5 patients per day for 5 days of the week. The patients must be consecutive and at least one day must be a weekend day;	*25 consecutive patients regardless of any specific day; excpet that it must include some weekend admissions;*	*If there are fewer than 25 admissions for a week, the total admissions for the week should be included in the sample.*
OR	OR	

Time was measured from the "decision to admit" to the arrival of the patient in the inpatient room. The destination could not be a "holding area" but had to be a real inpatient bed. The sample collection was done in real time. So, a data collection process had to be created by members of the team that were collecting the data. In this example, the collections had to be done weekly and summarized as the percentage of patients in the sample that achieved the goal for that week. Six weeks of data had to be collected and six data points placed on a run chart.

MEASUREMENT: RUN CHARTS

Why Use Run Charts?

There are many ways that data can be presented to tell the story of a project or improvement. Whether you use histograms, pie charts or run charts, the intention is the same: to gain new knowledge and to engage the audience, whether they are leaders, staff or customers. However, some graphical representations can be misleading.

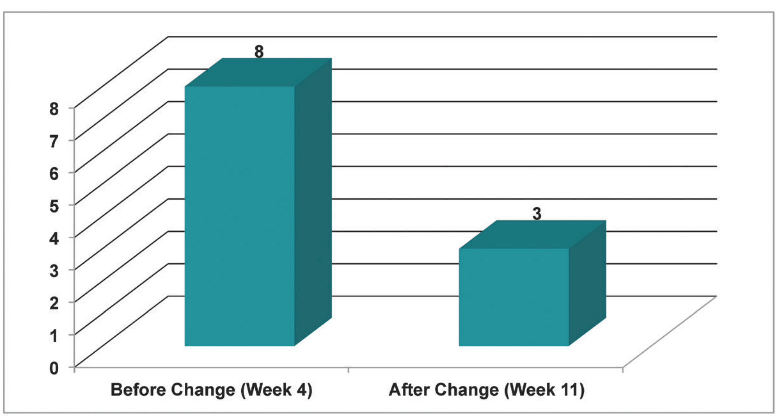
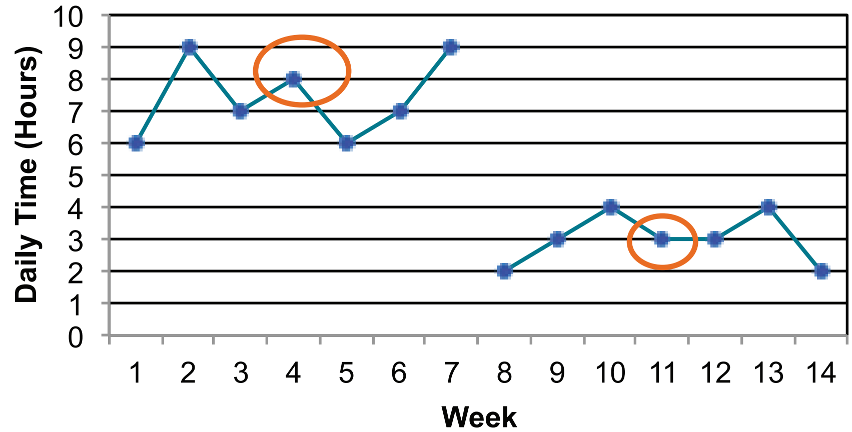


Figure 1

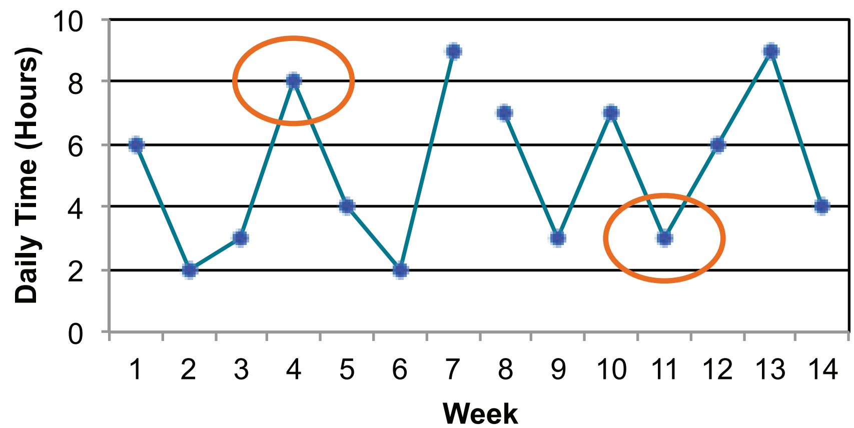
Figure 1 is an example of work by an improvement team formed in response to complaints from staff about delays in processing test results. This chart demonstrates the length of time (in hours) that it took for test results to be completed and received by staff. The graph demonstrates the changes measured at Week 4 and Week 11. During Week 4 (that is, four weeks after the team was formed), the team collected data to confirm or deny the complaints they were hearing. The data show that the turnaround time was eight hours, unacceptable by any standard. At Week 7, after the solution design process, the team tested a change. Measuring again during Week 11, they found that the turnaround time was now three hours. The reduction in cycle time from eight hours to three hours is significant and represents a 62.5% improvement.

Below are a number of scenarios that could have resulted in the above before-and-after bar graph. In each case, a run chart of the cycle-time for Weeks 1 to 14 is shown. The test results for Week 4 (cycle time of eight hours) and Week 11 (cycle time of three hours) are the same for all cases/scenarios. [4]

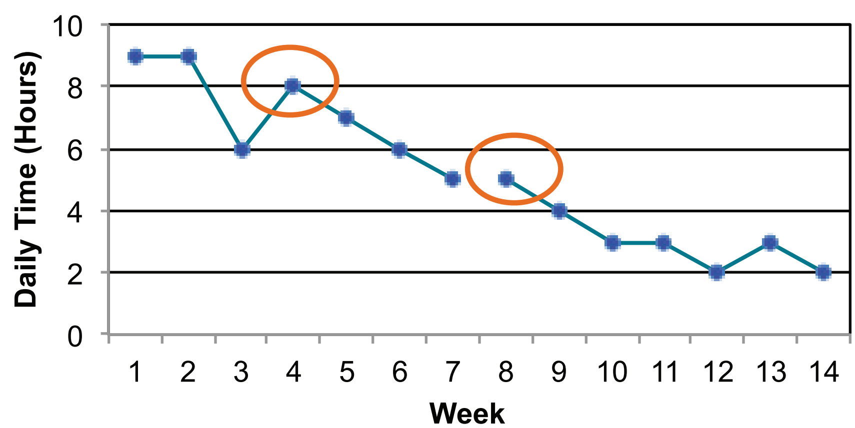
Scenario One *– this graph depicts 14 weeks of data, seven weeks prior to change and seven weeks post- change. This run chart would support the conclusion that the change resulted in meaningful improvement and that the process should become the new standard.*

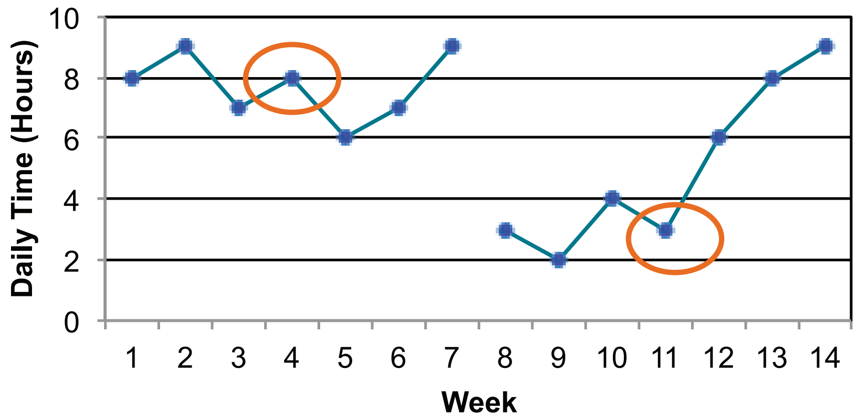


Scenario Two – *there is no obvious improvement after the change was made. The measures made during the test are typical results from a process that has a lot of week-to-week variation. The conclusion that can be drawn from a study of the run chart is that the change did not have any obvious impact on cycle time.*



Scenario Three – *It appears that the process steadily improved over the 14-week period. However, the rate of improvement did not change when the change was introduced in Week 7. Although the cycle time for the process certainly improved, there is no evidence that the change made any contribution to the steady improvement in the process.*





Scenario Four – *An initial improvement is observed after the change is made, but in the last three weeks the process seems to have returned to its pre-change cycle time. The results may be due to a Hawthorne effect, whereby an initial improvement is observed due to particular attention to the measures, but later, when focus on the change decreases, the cycle time reverts to the original process levels. The changes have not resulted in sustainable improvement. So the question here is, given just two numbers, can you be sure that the process that produced the second number is not the same as the process that produced the first number?*

A run chart can be used to display any measure over time and is very easy to develop, requiring not much more than a pencil and paper. Its simplicity makes the run chart a powerful tool and one of the most useful for understanding and communicating variation. Here are some of the reasons to depict your measures on a run chart:

1. Run charts can help you understand baseline performance and identify opportunities for improvement
2. They can help you determine if a change is an improvement
3. Once you have made an improvement, you can use the run chart to determine if you are sustaining the gains you have made
4. A run chart can be used to look at any type of measure over time. For example: costs, LOS (length of stay), counts, and percentages.

Run Charts to Engage Leadership and Staff

Run charts can also be a powerful tool for engaging leadership and staff. Without a clear picture of the actual outcomes, it is difficult to create a real desire for change or action around an issue. Quite often, staff are shocked when they are shown the performance of the organization over time and in a way that tells a story, which in turn can generate support for change. Also, it is difficult for leadership to create the business case for investing time and resources in an initiative without first understanding what the current system performance is and perhaps sharing this with a board or management team. Utilizing run charts to tell the quality story gets everyone on the same page and clears the path for the improvement to begin.

ANALYZING AND UNDERSTANDING RUN CHARTS

Variation

Driving to work is a form of variation that many of us experience. For example, your daily commute can take between 45 minutes and 60 minutes. There is 15 minutes variability for extra traffic or having to stop at all the stoplights along the route. This is common cause variation. Special cause variation is that snowstorm that causes our normal commute to take 120 minutes

One of the key strategies in quality improvement is to control variation. There are two types of variation: common cause and special cause.[5]

Common Cause variation is inherent in a system (process or product) over time, affecting everyone working in the system and affecting all outcomes of the system. A system that has *only* common cause variation is said to be stable, implying that the process is predictable within statistically established limits. Differences over time are due to chance rather than predictable influences on the system. Common cause does not mean *good* variation—it only means that the process is stable and predictable. For example, if a patient's systolic blood pressure is usually around 165 mmHg and is between 160 and 170 mmHg, this might be considered stable and predictable but it is also completely unacceptable.

Special Cause variation is not a usual part of the system (process or product), does not affect everyone, and arises because of specific circumstances which are not necessarily predictable. For example, special cause variation may be the impact of a flu outbreak on infection rates or the sustained impact of a targeted improvement activity to improve hand hygiene compliance.[6]

In the same way that common cause variation cannot be regarded as 'good' variation, special cause variation should not be viewed as 'bad' variation. You could have a special cause that represents a very good result (e.g., a low turnaround time), which you would want to emulate. Special cause merely means that the process is unstable and unpredictable.[7]

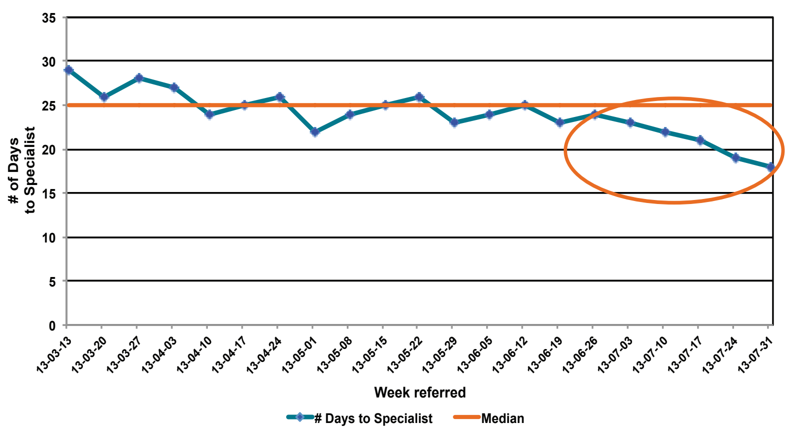
A system that has both common and special causes is called an unstable system. The variation may not be large but the variation from one time period to the next is unpredictable. Understanding the distinction between common and special causes of variation is essential to developing effective improvement strategies. When you become aware that there are special causes affecting a process or outcome measure, it is appropriate and usually economical to identify, learn from and take action based on the special cause. Often this action is to remove the special cause and make it difficult for it to occur again. Other times, the special cause produces a favourable outcome, in which case the appropriate action is to make it a permanent part of a process.

Because variation is normal and constant, data must be plotted over time to be useful, according to the authors of [The Improvement Guide: A Practical Approach to Enhancing Organizational Performance](). It is only by plotting data over enough time — both before and after a planned change is implemented — that you can judge whether the variation is random or forms a pattern that indicates that a meaningful change has occurred.

There are three signals of non-random change or special cause that you should look for on run charts. If you don't see evidence of one of these signals, then your data is exhibiting common cause variation. Finding one or more of these signals suggests that further analysis and interpretation by the team members is required in order to understand the causes or factors influencing the change. Keep in mind that not all common cause variation is good, and not all special cause variation is bad.

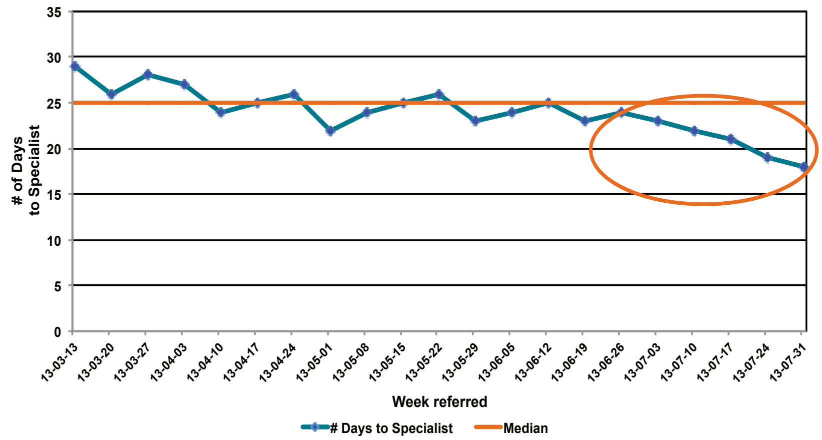
Signal 1: Shift

A shift signaling change is six or more consecutive points above or below the median. Values that fall directly on the median are not included in this count and neither break nor add to a shift.



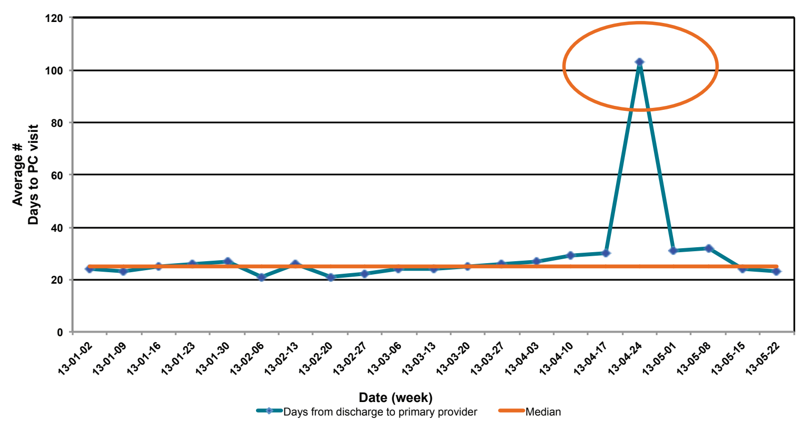
Signal 2: Trend

A shift signaling a trend is five or more consecutive points going up or down, excluding the starting point. Where the value of two or more consecutive points is the same, only include one in the count. For quality improvement, either there *is* a trend or there *is not*. Charts are not described as "trending."



Signal 3: Astronomical Point

An astronomical data point is one that is an obviously different value. Anyone studying the chart would agree that it is unusual. Every data set will have a highest point and a lowest point, but this does not necessarily make these points "astronomical." It is worth understanding the cause of the astronomical point but not necessarily to react to it. Understanding the reasons for this point will help the team emulate it if it is positive, and avoid or address it if it is negative.



As special causes are identified and removed or exploited, the process becomes stable. Deming identified several benefits of a stable process:[8]

- The process has an identity; its performance is predictable.
- Costs and quality are predictable.
- Productivity is at a maximum and costs at a minimum under the system.

The effect of changes in the process can be measured with greater speed and reliability.

- PDSA tests of change and more complex experiments can be used efficiently to identify changes that result in improvement.
- A stable process provides a sound argument for altering specifications that cannot be met economically.

1 IHI (2011). Science of Improvement: Establishing Measures. Retrieved from http://www.ihi.org/knowledge/Pages/HowtoImprove/ScienceofImprovementEstablishingMeasures.aspx

2 Ibid.

3 Kaplan, R.S. and Norton, P. (1997). *Balanced Scorecard: Translating Strategy into Action.* Boston: Harvard Business School Press, p. 28.

4 Provost, L.P. & Murray, S. (2011). *The Health Care Data Guide: Learning from Data for Improvement.* San Francisco: Jossey-Bass, pp. 10-12

5 Ibid, pp. 40-41

6 Ibid.

7 Ibid.

8 Deming, W.E. (1986) *Out of the Crisis.* Cambridge: MIT Press, p. 340

Health Quality Ontario
130 Bloor Street West, 10th Floor
Toronto, ON M5S 1N5
Tel: 416-323-6868 | 1-866-623-6868
Fax: 416-323-9261

ISBN 978-1-4606-1824-0 (PDF)

www.hqontario.ca