Using Run Charts for Healthcare Improvement – an Introduction

Steve Harrison

September 2012
VARIATION

Walter Shewart theorised as far back as in 1931 that the way to improve any process was to reduce the variation, and when feasible, to move the entire process in the desired direction. He distinguished two types of variation, *special cause* and *common cause* variation. A different approach to improve the process is needed depending on the type of variation. Run charts and control charts were developed as tools to distinguish one type of variation from another.

**Common cause variation** is random variation which can result from many factors inherent in a process. It is due to the regular rhythm of a process and produces processes that are stable or ‘in control’. One can make predictions within limits about a process that has only common cause variation.

**Special cause variation** is due to irregular or ‘special’ variation in the system that may be caused by something changing, or a special event that is not inherent in the process. When special causes are present, a process will be ‘out of control’ or unstable. Therefore one cannot make predictions on a process that exhibits special cause variation.

Neither type of variation is good or bad. However if a special cause variation is unplanned it is usually undesirable. However, when you want to change a process, a special cause ‘signal’ indicates that the change work has been effective. A process may exhibit common cause variation and be stable, but be totally unacceptable. For example a system showing post surgery mortality rate of 20% plus or minus 5% is a stable but unacceptable system.

**Improving the system.** When a process exhibits special cause variation the improvement effort should focus on understanding the special causes first rather than changing the whole system. Elimination of these special causes to make the system stable would be the focus of the effort. However with a system exhibiting common cause variation, the improvement strategy should be to focus on reducing the amount of variation and moving the average in the desired direction. This strategy involves implementing successive iterative changes and evaluating the effect of each change. This is the Plan Do Study Act or PDSA cycle. Run charts (and SPC charts) are useful in evaluating if the PDSA cycle has resulted in an improvement to the process. If a PDSA has had an impact on the system, there will be a special cause signal.

THE RUN CHART

To allow observers to distinguish common cause variation from special cause variation in their processes, Shewart developed a tool called the *control chart*. The prototype of a control chart is called a *run chart*. Both are diagnostic tools in determining the type of variation.

A run chart is very simple to construct and interpret. It can be used with any process and with any type of data: measurement data, count data,
percentages, ratios, and so on. On the other hand, there are several types of control charts. The choice among them is determined by the type of data (measurement data, count data, percentages, ratios, and so on). A run chart can be constructed with paper and pencil. It requires no statistics and is easily understood. Control charts are ordinarily constructed with the help of a computer software programme such as WinChart. They also are a little more complex to interpret. This paper will explain and demonstrate only run charts. References are given at the end of this paper which cover the more sensitive tool – the SPC or control chart.

Figure 1

What is a run chart?
Look at the run chart above (Figure 1). The run chart is a running record of a process over time. It is a different approach from merely comparing two aggregate measures, for example, 5% versus 4% falls rate. The units of time by which the measurements are made (in this instance, months) are located on the horizontal axis. The aspect of the process being measured (the number of falls) is on the vertical axis of the chart. (The horizontal axis is sometimes referred to as the “X” axis, and the vertical axis as the “Y” axis.)

The centreline of the run chart is the median. (Note - The centreline of a control chart will be the arithmetic mean of all the points.) Taking a sheet of paper and sliding it over the chart from the top down until half of the data points are above and half below the centreline can easily identify the median. In this case, with 18 points, there will be half (8) of the points above and half below the median, with 2 points on the median for this example.

What is a “Run”?
A run is defined as one or more consecutive data points on the same side of the median. A run could have a single point, or many points. In this
What are the tests for a Special Cause?
There are a number of run chart tests to identify a special cause that can be found in textbooks. Usually three tests are adequate.

Test #1: The presence of too much or too little variability. This is determined by analysis of the number of runs, and ascertaining if there are too few or too many runs based on the total number of observations. Table 1 displays a section of a probability chart developed by mathematicians. It is not necessary to understand how it was developed; just use it for reference purposes! Ordinarily, it is best to have at least 16 points in a run chart, excluding those on the median, to have adequate statistical power to identify a special cause. However, after 25 data points, additional data will not appreciatively increase the power of the tests.

In the run chart in Figure 2, there are 18 data points, and 10 runs. Table 1 shows that with 18 points, if there are less than six runs or more than 13 runs it is a signal that one or more special causes are present. In this case the number of runs falls within the lower and upper limits and therefore a special cause is not present.

Contrast this with Figure 3. The runs are once again circled and there are three on this example. The first is 7 points in length, the second twelve the third is five. Table 1 shows that with 24 data points, if there are less than eight runs; it is a ‘signal’ that one or more special causes are present.
Test #2: The presence of a shift in the process. A special cause exists if a run contains too many data points (that is, the run is too long). With 20 or more data points, a run of eight or more data points is “too long”. (With less than 20 points, a run of seven might also be considered “too long”.)

In Figure 2, the longest run is 3 points in length, so there is not a shift in the process. However in Figure 3, the longest run is 12 points in length, signalling the presence of a special cause, namely, a shift in the process.

Table 1 - Probability table for run charts.

<table>
<thead>
<tr>
<th>Number of Observations Excluding Points on the Median</th>
<th>Lower Limit for the Number of Runs</th>
<th>Upper Limit for the Number of Runs</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>15</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>16</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>17</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>18</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>19</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>20</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>21</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td>22</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>23</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>24</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td>25</td>
<td>9</td>
<td>17</td>
</tr>
<tr>
<td>26</td>
<td>9</td>
<td>18</td>
</tr>
<tr>
<td>27</td>
<td>9</td>
<td>19</td>
</tr>
<tr>
<td>28</td>
<td>10</td>
<td>19</td>
</tr>
<tr>
<td>29</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>30</td>
<td>11</td>
<td>21</td>
</tr>
</tbody>
</table>
**Test #3: The presence of a trend.** A trend is defined as an unusually long series of *consecutive* increases or decreases in the data. Experts disagree as to the exact number, but usually require that the number should be at least six or seven. The important lesson for healthcare managers is not to be "trend-happy", to be aware that three or four points are *not* considered a trend. For this test, count any points on the median, but ignore points that repeat the preceding value. In Figure 4 there are two trends. The first trend begins with clinic 7 and ends at clinic 12 and has 6 points. The second trend runs from clinic 15 to 22 and has 8 data points.

**Figure 4**

A run chart displaying too much or little variation (Test 1), A shift (Test 2) or a trend (Test 3) is defined as not being in statistical control. For improvement purposes the way to improve this process is to explain the special causes and try to eliminate the cause of this special cause variation. However if there is no special cause variation then the focus is to move the whole process.

Run charts are useful in understanding if improvement work is having a statistically significant impact on the measure you are trying to improve. Consider Figure 5 which shows patient waiting times in a Renal Outpatient Clinic. There are 10 runs so this is within the limits shown in Table 2 for 20 data points - there is not too much or too little variation. There are no shifts (8 points consecutively above or below the median) and no trends (6 points consecutively increasing or decreasing). So the process is showing common cause variation and is in statistical control.
The team then decide to introduce a new booking schedule and continue recording patient waiting times. The median line is fixed at 42.5. These results are then plotted on a new run chart with our baseline data on Figure 6.

Figure 6 shows the results of their PDSA. There has been a shift shown by the oval – with a run of 10 points below the median. The chart is no longer in control and therefore a statistically significant change has occurred.
change in schedule has resulted in a statistically significant reduction in waiting times. A new median line can be plotted – in this case it is 12.5 minutes.

CASE STUDY:
REDUCING DELAYS FOR REFERRAL TO TREATMENT – RAISED PSA

The situation

The medical director of Urology was concerned about what he felt were inordinate delays in the time between referral and treatment for the raised PSA pathway. He collected the average delay time per week for a period of 20 weeks (see Figure 7), presented them to the Service Improvement Team, and asked them to help remedy the problem. He felt an average waiting time of 31 days was unacceptable. He also felt that the process was continuing to deteriorate. He pointed to the fact that the delay time reached 40 days during week 14. In addition, during the last six weeks he felt there was a “negative trend” from 26 days to 35 days of wait time.

FIGURE 7

Questions

1. Does this process show common or special cause variation?
2. Is the current process stable and predictable?
3. What should the team’s strategy be to improve the process?
4. How will they know whether or not a planned change will be effective?
Analysis and interpretation

1. None of the three run chart tests indicated the presence of a special cause.
   - There were a total of 18 points, excluding the two points that were exactly on the median. There were 13 runs. The probability table (Table 1) showed that with 18 points, the upper limit for runs was exactly 13. Therefore, Test 1 showed only common cause variation.
   - The longest run was a run of four (from week five through eight), whereas a run of eight is required to indicate a shift in the process. Therefore, Test 2 indicated only common cause variation.
   - Test 3 requires at least six points consecutively ascending or descending for a statistical trend. However, in this process the longest series of consecutively ascending or descending points is only four (week’s three to six).
   - Some might observe the 40 days delay in week 14 and conclude this point is a special cause. However, it does not trip any of the tests for a special cause.

2. Because this run chart has 20 points and only common cause variation, the process is judged to be stable and predictable (median = 31 days delay). However, it is not functioning at an acceptable level.

3. The correct management approach is not to interpret the week with 40 days’ wait time or the four consecutive upward points as special causes. Investigating individual data points in a common cause system as though they were special causes will not be productive and may even be counterproductive. Instead of pulling patient charts and looking for someone to blame, the team should develop a plan to change the entire process. However, if a special cause had been present, the correct strategy would have been to investigate and eliminate it before deciding whether or not to change the process.

4. At the time the new process is implemented, the team should lock in (or “freeze”) the centreline (median) of the run chart, and then collect more data. If the next eight weeks show an average waiting time of less than 31 days, this would indicate a shift in the process, a special cause as determined by Test 2 of the run chart tests. The intervention would be a success. However, if the wait times were below the median for three or four weeks and then moved back above the median, or continued moving back and forth over the median, then the intervention was either not implemented properly, or the attempted solution was ineffective.

Further Reading (Including an introduction to SPC charts)

Carey (2003) – Improving Healthcare with Control Charts
Carey & Lloyd (2001) - Measuring Quality Improvement in Healthcare