

CHAPTER 8

Understanding Variation with Run Charts

Understanding variation from a conceptual point of view provides a solid foundation but it is only a start. If you are truly interested in immunizing yourself and those around you from numerical illiteracy, then you have to understand variation statistically as well as conceptually. This is where statistical process control (SPC) comes into the discussion. Although there are a number of tools and methods that can be grouped under the SPC term, two primary tools are essential for anyone interested in applying statistical thinking to daily work: the run chart and the Shewhart chart (a.k.a. control chart). These charts are used in order to:

- Make the process performance visible
- Determine if a change statistically demonstrates an improvement
- Determine if once an improvement has been observed it the improvement has been sustained (Provost and Murray, 2011, p. 85–86)

The remainder of this chapter addresses the construction, analysis, and interpretation

of run charts. Shewhart charts are discussed in the next chapter.

► What Is a Run Chart?

The run chart is a practical and easy-to-use statistical tool that is used to understand the inherent variation that lives in your data. Although there are many different types of Shewhart charts there is basically only one way to make a proper run chart. It is a plot of data over time with the unit of time (e.g., day, week, or month) always laid out along the horizontal or x axis and the indicator values (i.e., the data points) always plotted against the vertical or y axis. Any type of data (e.g., a count, a percentage, a rate, money, time, a score, an index, or days between events) can be plotted on a run chart. The data must be arranged in chronological order and the median is placed through the field of data points as a reference line.¹ **FIGURE 8-1** shows the basic elements of a run chart.

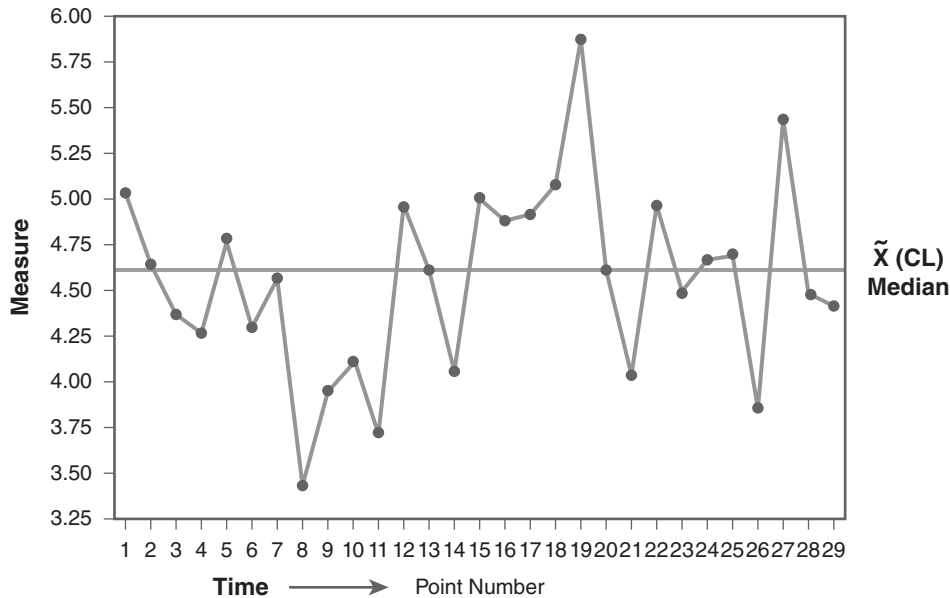


FIGURE 8-1 Elements of a run chart

► How Do I Construct a Run Chart?

Constructing a run chart is a very simple activity. It can be constructed by hand or with software. I have made many run charts on flipcharts during team meetings. I have even made them on a napkin or a piece of paper at a nursing station. The following steps should be followed if you wish to make a run chart by hand:

- Select the indicator of interest and make sure the data have been organized in time order. In terms of the number of data points you should have before constructing a run chart, there are mixed views. Purists would probably tell you that you should have 12–15 data points before making a run chart. This may pose a problem for a team, however, if they are tracking a measure that has no historical data. For healthcare applications I support the perspectives of my colleagues Lloyd Provost and Sandy Murray who write that “a run chart should begin when data for the first point to be plotted is available” (2011, 86).

The median can be placed on the run chart after four or five data points have been collected. You have to realize, however, that as you obtain more data the median will move a little. Typically it will not move as much as the mean when there is a small amount of data but it will fluctuate some. When you apply the run chart rules (described in the next section) you should have a minimum of 10 data points. This issue of how much data is needed to make a chart relates back to the discussion in the previous chapter on developing a data collection plan. Remember that run and Shewhart charts are designed to help you understand the variation in the current operation of a process. So if you are collecting data at the patient level (e.g., wait time for a patient to see the doctor) you will most likely have more than enough data to make a run chart in a day. But as you extend the time frame for your data collection frequency and move to days, weeks, or months you will have to wait a longer period of time to obtain sufficient data to make a run or Shewhart chart. Again, I would suggest steering away

- from even thinking about using quarterly data to make any of the charts.
- Lay out the horizontal (x axis) and vertical axis (y axis) and label them clearly. It is recommended that you extend the horizontal axis beyond the last data point you have collected so that future data can be plotted. On the y axis, which is the reference axis for the actual data values, make sure you extend this axis about $\pm 20\%$ more than the maximum or minimum data values in order to make sure the chart has ample space for future data values that may be greater than any that you have already collected. This will also provide enough extra space on the chart to accommodate a target or goal line. Also do not forget to make a clear title for the run chart that includes the name of the indicator, the unit, or facility involved and the dates for the period of data collection.
 - Plot the data in chronological order and make sure each data point is marked by a symbol of some sort (usually this would be a dot).
 - It is customary to connect the data points (i.e., the dots) with a line but note that there are some situations in which you would not connect the data points with a line.²
 - The centerline on a run chart is usually the median.³ Determining the median on a run chart is quite simple. Here is what you do:
 - First, you need to find the *median position*. This is the location in the distribution of data the median. This is analogous to the median's house number or where it lives in the distribution of data points. To find the median position use the formula $(n + 1)/2$. If you had 25 data points, for example, this would work out to be $(25 + 1) / 2 = 13$. In other words, the median lives at the 13th data position. Note that if you have an even number of data points, say 24, you will not get a whole number. If we put 24 into the median position equation we get $(24 + 1) / 2 = 12.5$. In this case, the median will reside between the 12th and 13th data points. In this case, you will need to average the two values found at the 12th and 13th positions to determine the median value.
 - Next, you need to find the *median value* (i.e., the number that actually resides at the median position). To do this you take a piece of paper or a ruler and place it at the top of the chart. Slide the paper or ruler down the chart and as you do you will reveal the data points in descending order. Because the y axis is calibrated from low to high, as you slide the paper down the page you are viewing the data in descending rank order, which is a requirement for calculating the median. As you slide the paper down the chart and reveal each data point place a checkmark next to it. Continue to reveal the dots until you reach the 13th data point (using our example where the median position = 13).
 - Now draw a horizontal line through the 13th data point and see where it intersects the y axis. The median value is the point on the y axis where the horizontal line ends. If we use the second example where the median position was 12.5 all you would need to do is draw the horizontal line equidistant between the 12th and 13th data points and see where that line intersects the y axis to find the median value.
 - To verify that you have done this correctly you can reverse the process by sliding the paper up the chart and you should end up at the same position as you did when you slid the paper down the chart and you should end up with the same median value.
 - A final point. The median position and the median value will usually not be the same. It is possible that this could occur but extremely rare. The median position is merely the location or place where you will find the median and the other is the actual numerical value of the median.

Sometimes explaining in words some of these statistical procedures can be more confusing than showing you how it is done. **FIGURE 8-2** provides a picture of how these steps are completed to discover the median value of 4.6. Note that this is the same chart as was shown in Figure 8-1 displaying the elements of a run chart. I keep referring to this same graph throughout this section so you can see the progression of constructing and interpreting a run chart. Once you are comfortable with this method for finding the median (shown symbolically as X with a ~ above it) you will actually be able to compute the median faster by hand than entering all the data into the computer and using software to produce the median value.

► How Do I Analyze a Run Chart?

Once you have laid out a run chart, the first question to be answered is, “What is a run?”

A run is defined as one or more consecutive data points on the same side of the median (Lloyd, 2010). **FIGURE 8-3** shows how runs are determined. The first thing to do is to identify the data points that fall exactly on the median. These data points need to be identified but they are ignored when determining the number of runs. You can count the number of runs by (1) drawing a circle around each run and counting the number of circles you have drawn or (2) count the number of times the sequence of data points (the line on the run chart) crosses the median and add 1 to this count. If you have done this correctly, the number of circles you have drawn around the clusters of data points (i.e., the runs) should match the number you get when you count the number of times the data line crosses the median plus 1.⁴

In Figure 8-3, there are two data points that fall exactly on the median value of 4.6. These two data points are highlighted by placing a square around each one to remind us to

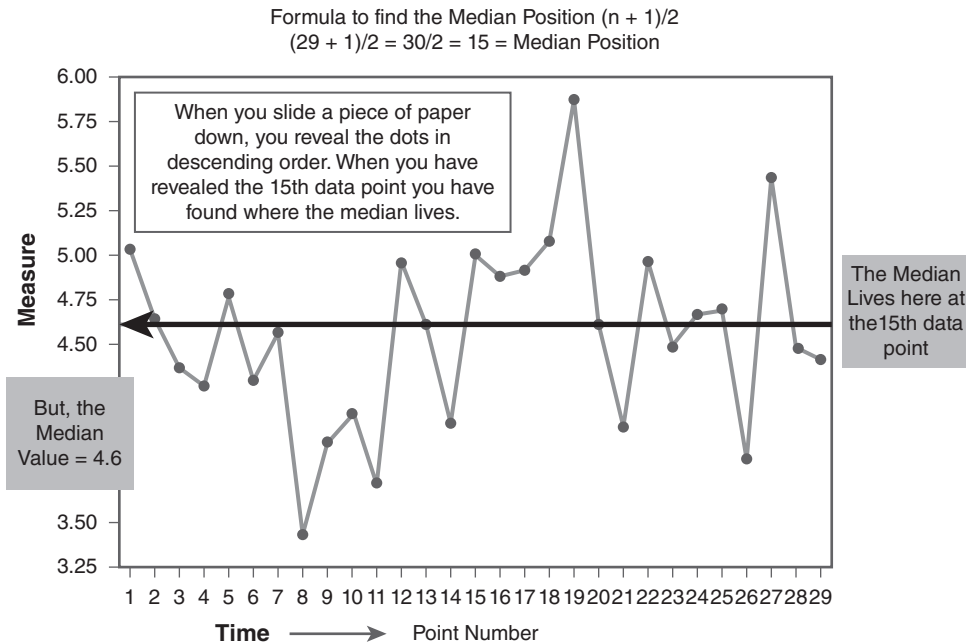


FIGURE 8-2 Locating the median on a run chart

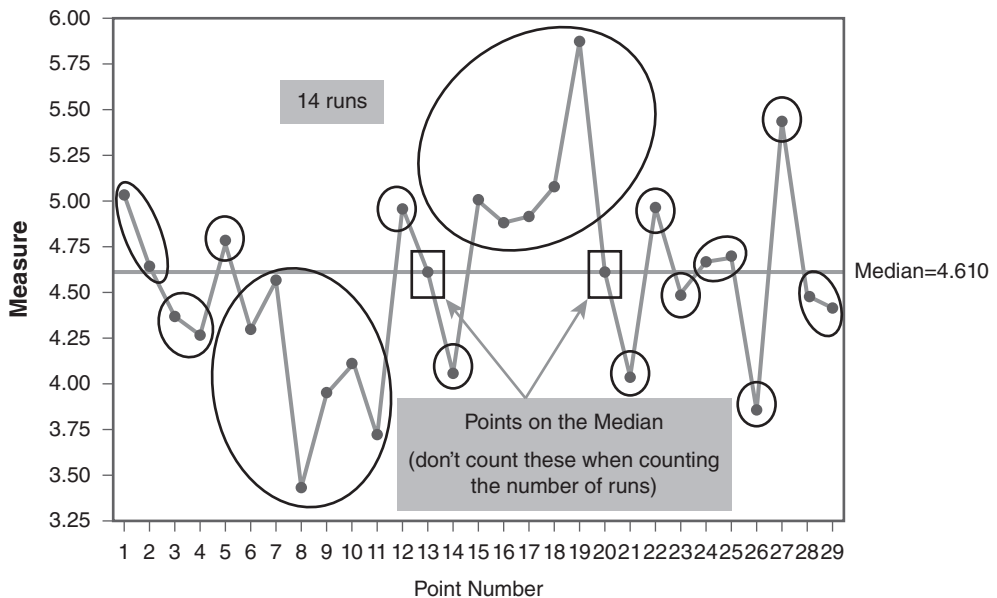


FIGURE 8-3 Determining the number of runs on run chart

exclude them from the identifying the number of runs. Circles have been drawn around the runs in Figure 8-3. Fourteen runs are present on this chart. If you count the number of times the data line crosses the median you will end up with 13 crosses. If you add 1 to this number you end up with 14, which is the same as the number of circles highlighting the runs. The first run contains two data points, then the next run of two data points is below the median line. Then the next run of one data point goes back above the median. Remember that a run is defined as one or more consecutive points on the same side of the median, so you can have a run with one data point in it. The other technical point to remember when counting runs is that because you are ignoring data points that fall on the median, these data points do not get included in the determination of a run. You can see this in Figure 8-3. Notice that the circles do not include the data points on the median (the median data points are identified by the two boxes).

When people are just starting out with run charts, they have a tendency to (1) include the data point(s) on the median in the run, or (2) count the data before the point on the median as one run and then count the data after the point on the median as a second run. For example look at **FIGURE 8-4**. The chart has six runs shown by the circles. If you count the number of times the data line crosses the median you get 5 plus 1 equals 6 runs. So either way you get the same result. The first run in Figure 8-4 contains six data points. It does not contain seven data points, which would include the median data point, and it does not reflect two separate runs broken by the second data point, which falls on the median. The same logic is the same for the fourth run, which contains five data points. Notice that there is a data point on the median but it is not part of the run. This point of ignoring points on the median when determining a run is probably the only real challenge when determining the number of runs but with a little practice you will master this step quickly.

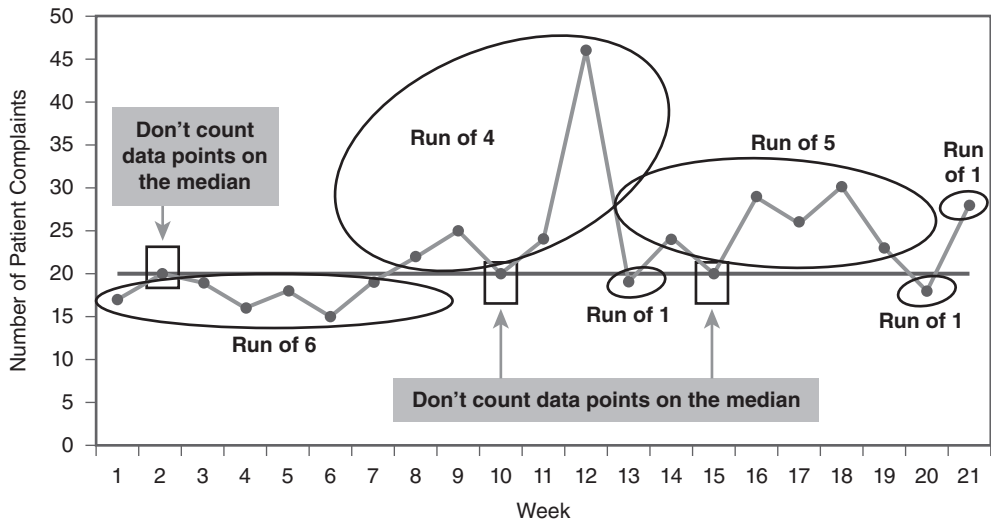


FIGURE 8-4 Determining the number of runs on a patient complaints run chart (6 runs)

Once we have identified the number of runs on the chart, the next question that most people ask is “So what? I’ve made this nice little line graph and circled some clusters of data, but why did I do it?” What can we learn from the number of runs? The answer should be obvious. The number of runs allows you to apply statistical rules that will help you identify the type of variation that lives within your data. A technical caveat needs to be noted at this point. Earlier we discussed the types of variation and pointed out that variation can be classified as either common cause or special cause. These terms technically should be reserved for use with Shewhart charts. Because run charts are not as precise as Shewhart charts (think x-ray versus an MRI) the proper terms for describing variation on a run chart are random and nonrandom patterns rather than common and special causes of variation. Now, I know some of you are thinking, “Picky, picky, picky.” You will hear many people using the terms common and special to describe the variation patterns on a run chart. You will have to decide for yourself if it is worth pointing out to these folks that the terms common and special should be applied only to variation analysis on Shewhart charts and not on run charts. But, at least you are now aware of this technical point.

It is important to realize that there are a variety of run chart rules that different writers have offered over the years to identify nonrandom signals on a run chart. Some of these rules, although having the same name, for example a “trend,” will use different decision criteria to decide if a trend exists or not. This is another important aspect of run and Shewhart chart analysis. So, do not get too concerned if when you present a chart and start to explain how you have analyzed it someone says, “But that is not what I was taught was a trend.” Life is full of options and even among statisticians there can be considerable variation.⁵

At the Institute for Healthcare Improvement (IHI), we have reached agreement on four run chart rules that we have found to be most relevant to healthcare applications:

- Rule 1: A shift in the data
- Rule 2: A trend in the data
- Rule 3: Too many or too few runs in the data
- Rule 4: An astronomical data point

Rule 1: A Shift in the Data

When do you know if a process has really moved to a new level of performance? If you suffer from numerical illiteracy, you will probably think that one or two data points that are “better” (however

defined) constitute a shift to an improved level. When looking at data on a run chart, however, the key to defining a shift in the process is when you have a run that contains six or more consecutive data points on one side of the median or the other (Lloyd, 2010). According to Provost and Murray (2011, 77) the probability of this occurring by chance alone when there has been no real change made to the process is less than 5% chance. In Figure 8-3, a shift is observed in the fourth run where six data points are in a run that is below the median. The question is what were the conditions affecting the process that caused the data to stay below the median for six points in a row? The run chart rule does not provide an answer but merely points you to a period when a nonrandom pattern was detected. In Figure 8-4, the process starts out with a shift of six data points in a row below the median. Then a random pattern emerges for the rest of the chart. Often when a run chart starts out with a shift below the median it can be because the team was still learning how to collect the data and properly apply the operational definitions for the indicators. After a period of learning how to apply the operational definition and data collection plan, however, the true performance of the process is observed. Again, the detection of a nonrandom rule does not provide the answers but should lead the team to raise question.

Rule 2: A Trend in the Data

This is probably one of the more difficult rules to apply because of all the popular notions about what constitutes a “trend.” People love to find a trend. This issue was discussed earlier in this chapter and is not elaborated further in this section. The key point to remember, however, is that at the IHI we define a trend statistically as five or more consecutive data points constantly going up or constantly going down (Lloyd, 2010; Provost and Murray, 2011). This choice, which is at odds with other decision rules on trends (see for example Pyzdek, 1990; Western Electric Company, 1985) is based on our practical experience of applying run chart rules to healthcare

data and the fact that more stringent criteria for a trend (e.g., six or seven constantly increasing data points for a trend) require more data that many healthcare teams do not have the luxury of collecting, especially when they are just starting their improvement efforts. So, the decision to use five constantly increasing or decreasing data points as a trend is grounded not only in statistical theory (i.e., the probability of getting five data points that constantly increase or decrease is grounded in the same probability as getting five heads in a row on a coin flip or .031 probability) but also practical experience. Two final points to remember when deciding whether a trend exists: (1) the median does not come into consideration when determining whether a trend exists so a data point that falls on the median can be considered as part of a trend, and (2) ignore points that repeat the previous value. **FIGURE 8-5A** shows a run chart with 29 data points with 1 on the median (marked by a square) and 13 runs. Do you see a trend in this chart? Actually there are two downward trends highlighted in **FIGURE 8-5B**. Notice how both of the trends cross over the median. Remember that a trend is not dependent on the median. Rules 1 and 3, however, do require the median to make a determination as to whether nonrandom variation is present in the run chart.

Rule 3: Too Many or Too Few Runs in the Data

Another nonrandom pattern is defined as too much or too little variation in a dataset. Like Rule 1, Rule 3 is based on probability theory. When applying this rule, the first thing that you need to do is to determine the number of “useful observations” in your dataset. As stated earlier, this is determined by subtracting the number of data points on the median from the total number of data points. Once you have determined the number of useful observations in your dataset, you refer to a table (**TABLE 8-1**) to identify the minimum and maximum number of runs that should be observed statistically for each respective amount of data that is defined as useful observations.

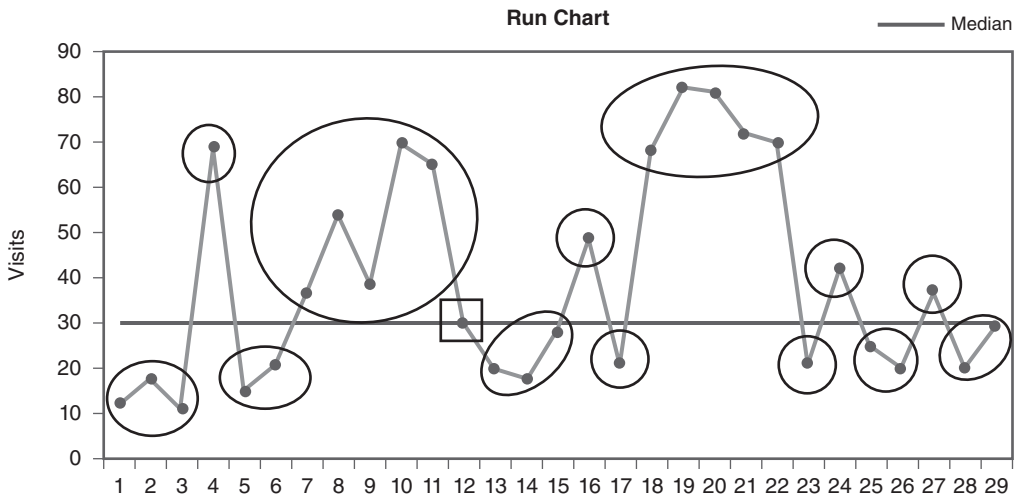


FIGURE 8-5A Run chart with 13 runs

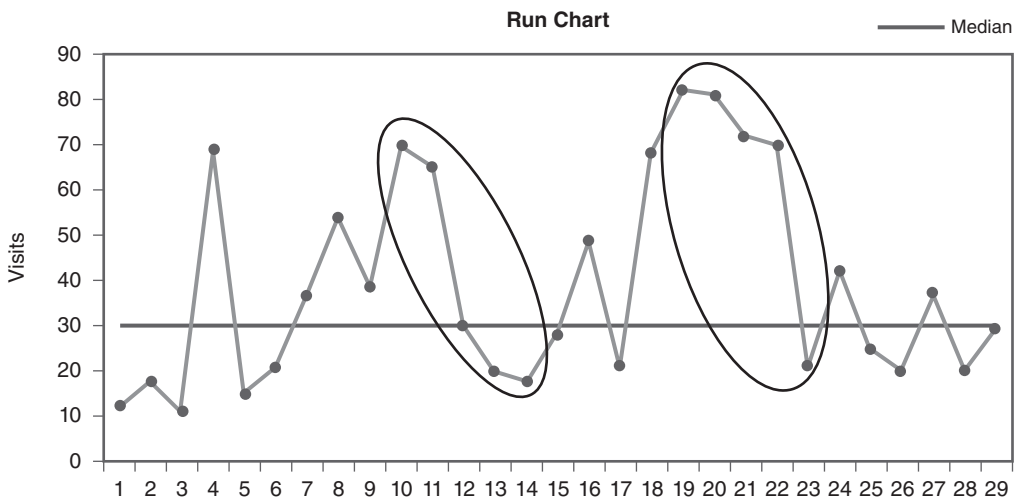


FIGURE 8-5B Run chart with 13 runs and two downward trends

The table most commonly used is one developed in 1943 by Swed and Eisenhart. If the number of runs on your chart falls between this minimum and maximum number of runs then you do not have a nonrandom pattern. For example, if you had 23 useful observations Table 8-1 indicates that the minimum number of runs should be 7 and the maximum number of runs should be 17. If your run chart had less than 7 (i.e., too few runs) or more than 17 runs (i.e., too many runs)

this would be a statistical signal that your data exhibit a nonrandom pattern.

To properly use Table 8-1 you:

- First calculating the number of “useful observations” in your dataset. This is done by subtracting the number of data points on the median from the total number of data points.
- Then, find this number in the first column. The lower number of expected runs for this

TABLE 8-1 Table to determine whether there are too many or too few runs on a run chart

Number of Useful Observations	Lower Number of Expected Runs	Upper Number of Expected Runs
10	3	9
11	3	10
12	3	11
13	4	11
14	4	12
15	5	12
16	5	13
17	5	13
18	6	14
19	6	15
20	6	16
21	7	16
22	7	17
23	7	17
24	8	18
25	8	18
26	9	19
27	10	19
28	10	20

(continues)

TABLE 8-1 Table to determine whether there are too many or too few runs on a run chart*(continued)*

Number of Useful Observations	Lower Number of Expected Runs	Upper Number of Expected Runs
29	10	20
30	11	21
31	11	22
32	11	23
33	12	23
34	12	24
35	12	24
36	13	25
37	13	25
38	14	26
39	14	26
40	15	27

Excerpted from: Swed, F. and Eisenhart, C. (1943) "Tables for Testing Randomness of Grouping in a Sequence of Alternatives." *Annals of Mathematical Statistics*. Vol. XIV, pp. 66–87, Tables II and III. This is a segment of a longer table that goes from 10 to 60 useful observation.

number of useful observations is found in the second column.

- The upper number of expected runs can be found in the third column. If the number of runs in your data falls below the lower number of expected runs or above the upper number of expected runs then this is a statistical signal of a nonrandom pattern.

FIGURE 8-6 shows a run chart with too many runs. The chart has 25 data points with 1 data point on the median so there are 24 useful data

points. Table 8-1 indicates that for 24 useful data points we should find 8 to 18 runs if the chart displays only random variation. Figure 8-6 has 21 runs (20 crossings of the median plus 1 = 21 runs), which exceeds the upper number of expected runs of 18 runs thus signaling the presence of a nonrandom pattern. This is a signal that the number of home care visits reflects a nonrandom pattern that is unstable and therefore unpredictable. The team needs to investigate why the number of home care visits keeps fluctuating so much.

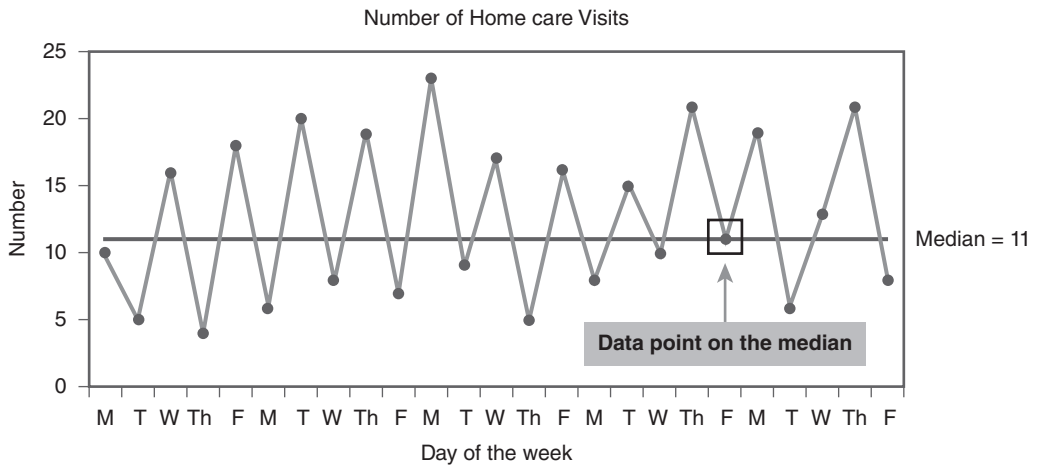


FIGURE 8-6 An example of too many runs on a run chart

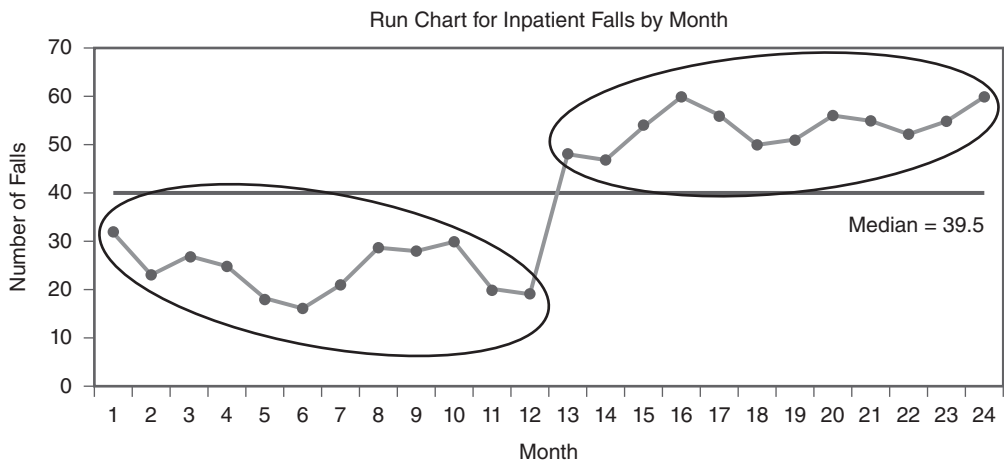


FIGURE 8-7 An example of too few runs on a run chart

The other aspect of Rule 3 (i.e., too few runs) is shown in **FIGURE 8-7**. This run chart contains 24 data points with no data points on the median, which gives us 24 useful observations. If we look up 24 useful observations in Table 8-1 we see that if the data exhibited just random variation we should observe between 8 and 18 runs but we detect only 2 runs. Did you also notice that Rule 1 is present in Figure 8-7? There are two shifts in the data. Both contain runs with 12 data points. The occurrence of Rule 1 and Rule 3 will often be observed together. The reason is that

if you have several runs on a chart that contain six or more data points indicating shifts in the data, then this will reduce the opportunity to have more runs on the chart and increase the probability of having too few runs. In the case of Figure 8-7, half the data are hanging on one side of the chart and the other half of the data have shifted (for some unknown reason at this point) above the median. The challenge for the team is to determine why the number of inpatient falls has shifted upward to a new (and probably less desirable) level of performance. Did they

institute a change between month 12 and 13? As an improvement advisor to this team I would ask them to annotate the run chart and show where they made changes in the process so we could see if in fact their intervention(s) have caused the shift. If their actions do not coincide with the shift then we would need to explore other possible cause for the shift.

Rule 4: An Astronomical Data Point

This is not a statistical rule but an observational rule that serves as a filter to guide further investigation and dialogue. I refer to this as the interocular test of significance. An astronomical data point is *one data point* that is dramatically different from the rest. Look at **FIGURE 8-8A** for an example of an astronomical data point. Data point 24 on the far right of the run chart clearly looks like it is very different from the field of data points lower on the chart. All you have to do is look at the chart and consensus would most likely be that data point 24 is different. The key to applying this rule is that it relates to only one data point. In every dataset, you will have a high data point and a low data point or possibly several high and several low points. These are not astronomical data points. In fact, when you

have a high and a low data point they basically balance each other out.⁶ We are looking for one point that is egregiously different from the rest. Now what do we do with this data point? You have several options:

1. First, I would investigate whether this data point is based on the same operational definition as the other data points. Someone new to the team who was collecting the data might have and included (or excluded) observations that were not included (or excluded) in the rest of the data. It is also possible that the data collection process was different for data point 24. So when you see an extreme data point before jumping to conclusions check the operational definition and data collection plan that was used during the week in question.
2. If you are satisfied that a different operational definition and/or the data collection plan used during week 24 was not different from that used for the rest of the data, then it is time to place all 25 weeks of data on a Shewhart chart and see whether the data for week 24 exceeded the upper control limit (UCL). **FIGURE 8-8B**

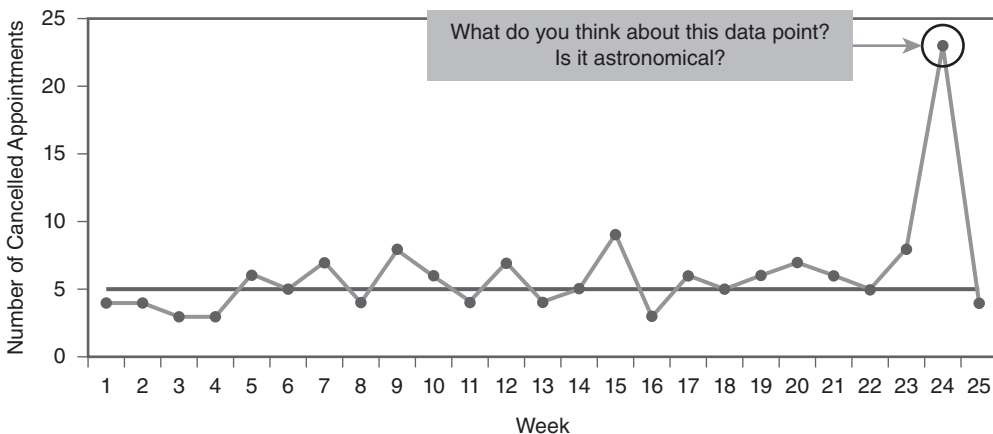


FIGURE 8-8A An astronomical data point

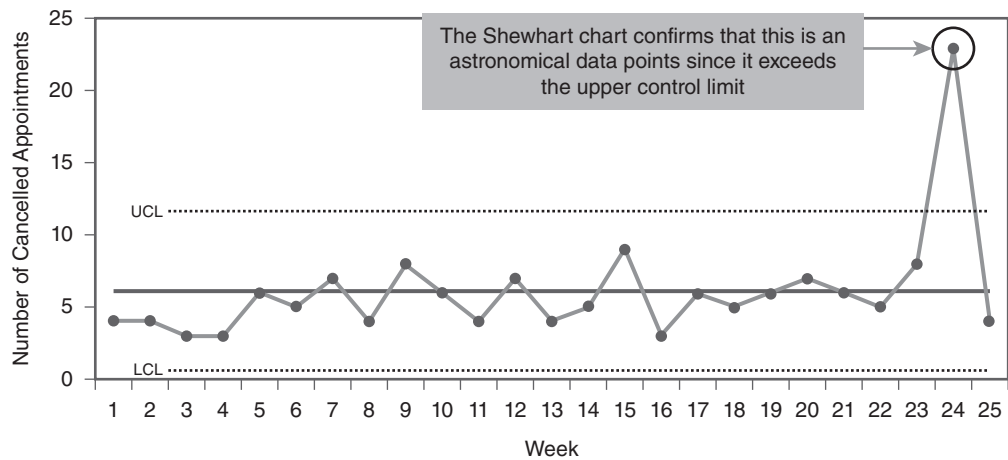


FIGURE 8-8B Verifying an astronomical data point on a Shewhart chart

shows how this would be done. Placing the data on a Shewhart chart will provide clear confirmation as to whether this data point was astronomical or not. More will be said on the use of control limits in the next chapter.

3. Finally, the team needs to conduct a root cause analysis and determine what the conditions were during week 24 that led to a much higher number of cancelled appointments. If they do

not investigate this data point and take steps to extricate the factors that caused this astronomical data point to occur then they can expect to see the point pop up at some point in the future. No one can predict exactly when this will happen again. But because the conditions that caused it initially would still be inherent in the process or system (Reason, 2003) you can count on it happening at some point in the future.

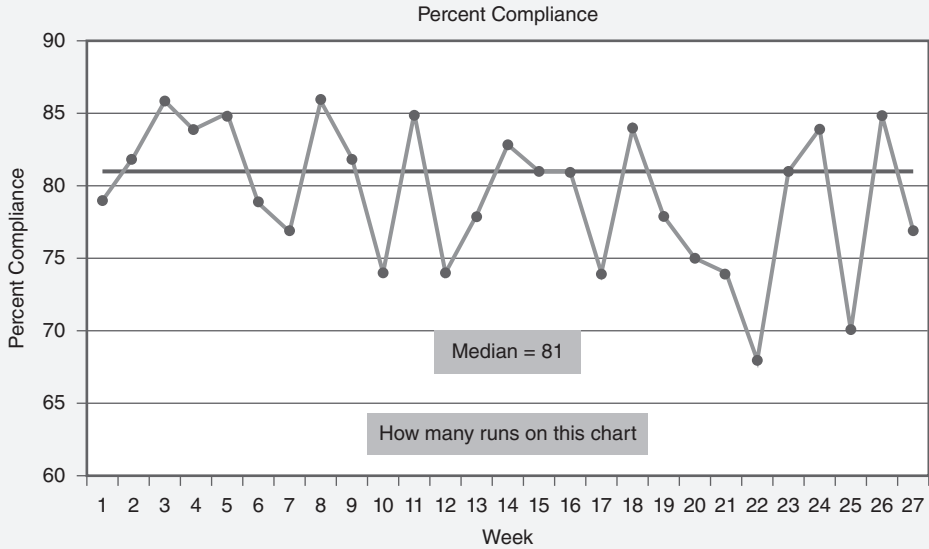
CASE STUDY #1: Hand Hygiene Compliance

As improvement advisor to a team trying to improve hand hygiene compliance you told them that at the next meeting you would show them how to make a run chart and determine if their recent interventions have made a difference in the key indicator, percentage of compliance with proper hand hygiene. The indicator is defined as the percentage of compliance with proper hand hygiene by week where the numerator is the total number of properly completed hand washings and the denominator is the total number of hand washing observations performed in the week.

The data the team collected each week were obtained through observations of a stratified random sample of staff on three different units. The data along with the initial run chart you created on your laptop in the meeting are shown in **FIGURE 8-9**.

(continues)

CASE STUDY #1: Hand Hygiene Compliance (continued)



Week	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
Percent Compliance	79	82	86	84	85	79	77	86	82	74	85	74	78	83	81	81	74	84	78	75	74	68	81	84	70	85	77

FIGURE 8-9 Hand hygiene compliance data for 27 weeks and initial run chart

Even though you made the run chart on your laptop using SPC software, you provide all the team members with a paper copy of the chart so you can:

- Walk the team through the calculation of the median and make sure they understand how it was obtained using the formula to find the median position $[(n + 1)/2]$. In this case, 27 data points plus 1 = 28 divided by 2 = 14 which is the median position or the location of the median in the array of data.
- Then you have them all slide a piece of paper down the chart to reveal the data points in descending order.
- Because the formula indicated that the median position is 14 you have them stop revealing the data points when they arrive at the 14th point. They then discover that there are actually three data points at this location.
- You then have them draw a line through the points at the 14th position and have it intersect the vertical (y axis). This line points to the median value of 81. You have the team members confirm this by referring to the data in Figure 8-1 where they find that there are in fact three data points at the value of 81% compliance.
- The next question for the team is, "How many runs are there on this chart?"
- You ask them all to determine the number of runs and when they have a number to shout it out. Someone says "14" then another says "15" and after a period of contemplation another team

(continues)

CASE STUDY #1: Hand Hygiene Compliance

(continued)

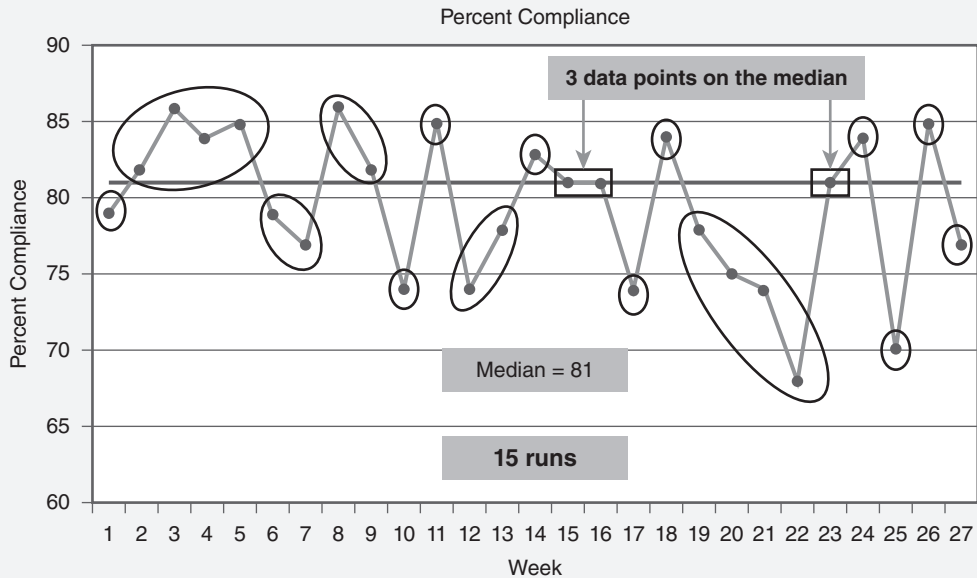


FIGURE 8-10 Number of runs for the hand hygiene compliance run chart

member says “16.” So now the learning begins. You describe the definition of a run, which you had covered at the last meeting, then show them the two methods for determining the number of runs (i.e., draw circles around the clusters of dots on either side of the median or count the number of times the data line crosses the median line and add 1). Both methods should produce the same number of runs. After a little reflection the team arrives at consensus: 15 runs as shown in **FIGURE 8-10**.

- Now that the team members know the number of runs on the chart the next step is to have the team analyze the run chart and determine whether any of the run chart rules are detected. Note that as the improvement advisor to the team your job is not to analyze the chart for them. The data are their data not yours. Your job is to provide guidance not answers.
- You review the four run chart rules with the team (that you printed on the back of the chart you gave them) and ask them to interpret the chart. **FIGURE 8-11** shows the result of the analysis. A downward trend was detected from weeks 18–22, then random variation was observed after week 22. Even though the team had put in place a hand washing campaign during weeks 16 and 17 they realized that this was prior to the start of summer vacations and the rotation of the resident doctors. So although the results were not what the team expected they did gain great insights about factors affecting their improvement efforts. Now, the next question is, “What are you going to do to make sure the process works reliably during the next vacation period and when residents come on the scene?”

CASE STUDY #1: Hand Hygiene Compliance

(continued)

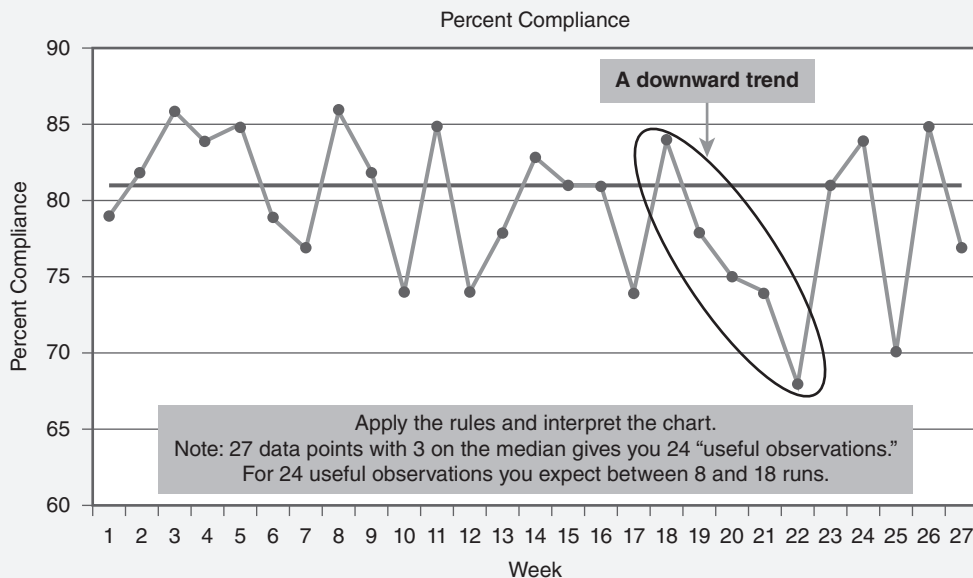


FIGURE 8-11 Interpreting the hand hygiene run chart

► A Few Closing Thoughts on Using Run Charts

As you gain more knowledge of and experience with run charts there are a few additional things to think about. Specifically the following questions are the ones I get asked most often:

- Is there a minimum and maximum amount of data I can put on a run chart?
- When do I change the median on a run chart?
- Why do I need to annotate the run chart?

Is There a Minimum and Maximum Amount of Data I Can Put on a Run Chart?

At the beginning of this section I mentioned that as soon as you start collecting data you should

start to plot it. At first this can just be a simple line graph. When you get four or five data points you can place an initial median on these data to see how the data are centering around a common point but I find it useful to refer to these early medians as “trail medians” because the median may change until you obtain sufficient data that you have a reasonably stable distribution. This stability will start to occur with a median when 10 or more data points are on the chart (Provost and Murray, 2011). At the other end of the spectrum is the question of the maximum amount of data that should be placed on a run chart. You can easily place a lot of data on a run chart especially if you use a computer. In general, however, once you get upwards of 40–50 data points you have exceeded the practical design of what a run chart is trying to convey.

Two issues will influence the amount of data you should consider placing on the chart.

First is the issue of current versus old data. Remember that the primary use of a run chart (or Shewhart chart) is to help you understand current performance. So when you place data that go back 2, 3, 4, or more years on a chart these are historical data not real-time data. Wheeler and Chambers refer to charts like this as report card charts. They point out that although the chart may be valid and constructed properly it is a “weak use of the charts” (1992, p. 18). The second issue that will influence the amount of data you place on a chart depends on how you have organized the data (e.g., by patient, day, week, or month). How you have organized your data, referred to as the subgroup, will provide different amounts of data for the chart. For example, if you are tracking patient wait time to see the doctor in a family practice clinic that sees approximately 50 patients a day you will end up with over 200 data points in a week if you recorded every patient’s wait time to see the doctor. This is way too much data. If you make the subgroup day instead of every patient, however, and select a sample of patients each day (e.g., 5–10) then you will have a more manageable amount of data to place on a chart.

When Do I Change the Median on a Run Chart?

The answer to this question is essential in order to decide whether the change a team has made to a process or the system actually made a difference. Here are the steps to follow in order to answer this question:

- Establish a baseline for the indicator of interest. Ideally it would be nice to have a solid baseline that established the performance of the indicator prior to any intervention. Again, the number of data points in this baseline will depend on how you have organized your subgroups (i.e., the labels on the horizontal axis of your run chart) and your data collection plan. The following are not strict recommendations based on hard and fast statistical rules but rather guidelines

I have used with teams that have proven to be practical yet useful. You will have to discuss with your team what seems to be reasonable for your work.

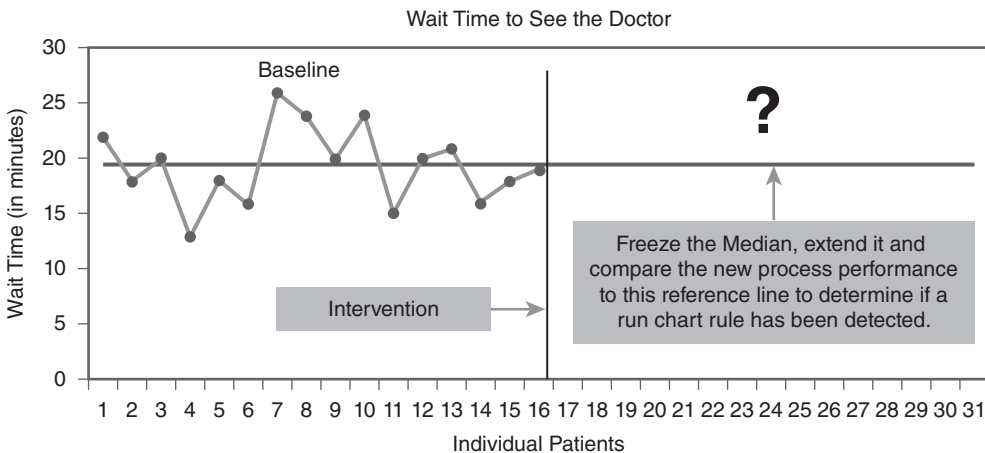
- If you have organized your data into monthly subgroups it would be desirable to have approximately 10–12 months as the baseline. However, you may need to make compromises. For example, when I was setting up the measurement system for our work with the Scottish Patient Safety Program (SPSP), I made a practical decision because many of the indicators had not been collected before at the 42 hospitals. So, we settled on collecting the first 6 months of data for newly developed indicators and calling this the baseline. Not ideal but practical.
- If you have sufficient data to organize your data into weeks as the subgroup, then having a baseline of 10–12 weeks would provide an acceptable baseline (e.g., the percentage of newly admitted inpatients properly assessed for pressure ulcers each week).
- If you have organized your data into daily subgroups then 12–15 days of data could serve as a baseline (e.g., the number of food trays produced each day).
- If you have organized your data by individual patients, tests, surgeries, or some other discrete unit that is essentially an n of 1, then I would build a baseline with roughly 25–30 individual occurrences (e.g., the wait time for each patient to see the doctor at the family practice clinic or the turnaround time in minutes for a stat medication order).
- Once you have a baseline, calculate the median for this block of data. Then you “freeze” the median for the baseline and extend it into the future as a reference marker. Ideally the baseline period is the time when the process was functioning without any improvement work occurring. At the end of the baseline

period you would draw a vertical line on the run chart to show that this is the point at which the team began to test new ideas that they believed would improve the performance of the indicator being tracked. The vertical line on the chart provides a visual demarcation point between the old way and the improvement journey. The question is, what will happen to the indicator after the team tested new ideas for improvement. The new data will be plotted against the baseline median. The run chart rules will then be applied to the new data using the baseline median as the reference point. This is shown in **FIGURE 8-12A**.

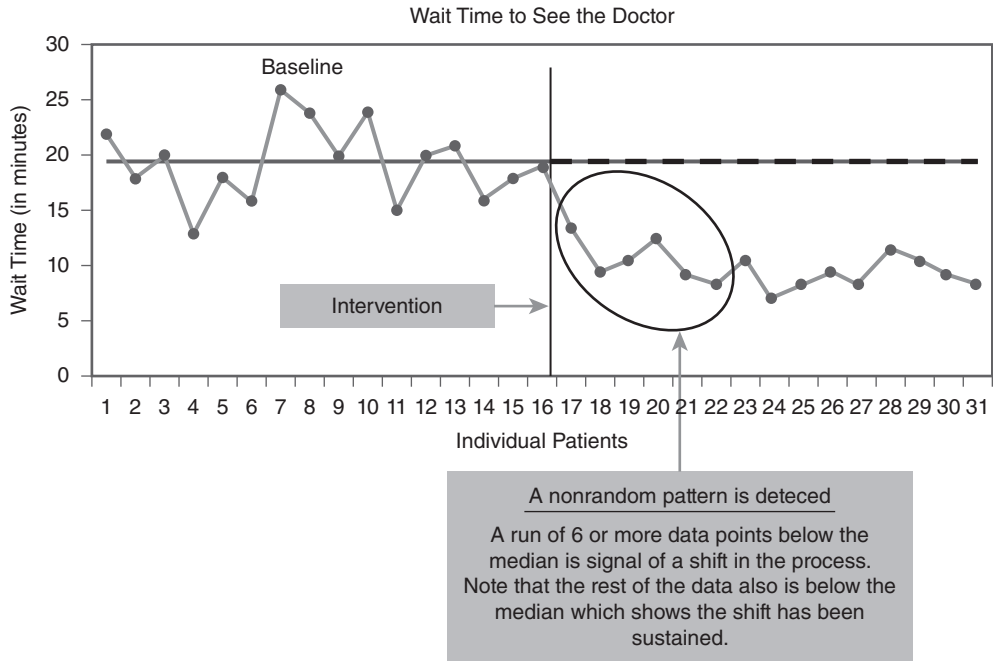
- **FIGURE 8-12B** shows how the new data after the intervention performed against the baseline median that was extended. Not only did six data points stay below the baseline median indicating a shift in the data but the remainder of the new data also stayed below the baseline median. This not only indicates that there has been a downward shift in the wait times but also that this shift has been sustained.
- **FIGURE 8-12C** shows the final step in calculating a second median. Once the nonrandom pattern (Rule 1: a shift) has been detected it is now appropriate to create a new median

for the new data. In Figure 8-12c, we see that not only has the median shifted downward but also the variation in the wait time to see the doctors has been reduced (i.e., the dots are all closer together). This actually indicates that two improvements in the process have occurred: (1) a shift in the median in the desired direction and (2) a reduction in the variation.

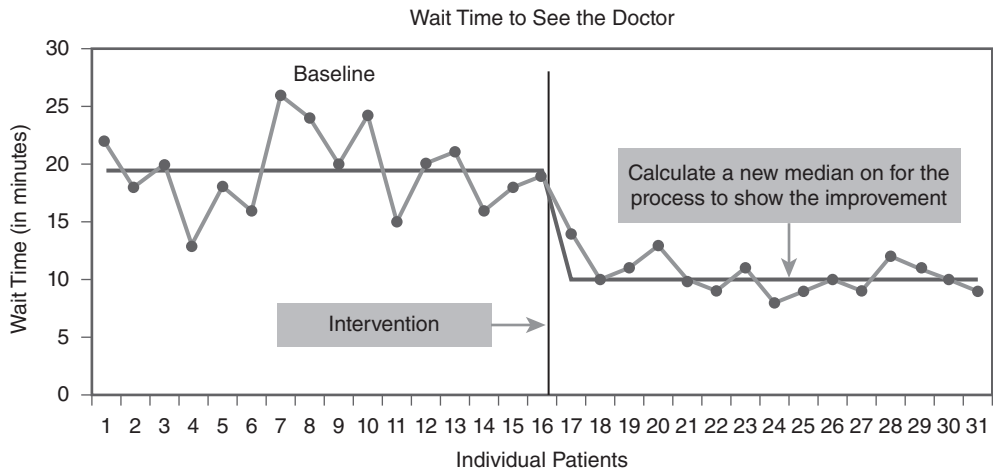
- If the team planned further improvements, this sequence of freezing and extending the second median, making the next intervention, and then plotting the new data against the second frozen median would be the strategy. If after making the next improvement intervention the data did not demonstrate a shift or trend then we would conclude that the intervention did not have the expected impact on the process. The team would then need to rethink its strategy. Was the idea the wrong idea? Did we have the wrong theory about how this idea would affect the process? Did we collect the data in the same way we did previously? All these questions would be part of the team's learning based on the results they observed.
- In summary, the sequence for creating a new median is to establish a baseline, create the initial baseline median, freeze it, and



FIGURES 8-12A Freezing the median on a run chart



FIGURES 8-12B Plotting the new data against the baseline median



FIGURES 8-12C Creating a second median on new level of performance

then extend it into the future. Next, put in place the new idea that the team believes will make an improvement in performance. Plot new data against the frozen median to see if a run chart rule is detected. If it

is, then start the sequence all over. On the other hand, if a run chart rule is not evident then the team needs to evaluate why the improvement idea it introduced did not have the expected impact.

Why Do I Need to Annotate the Run Chart?

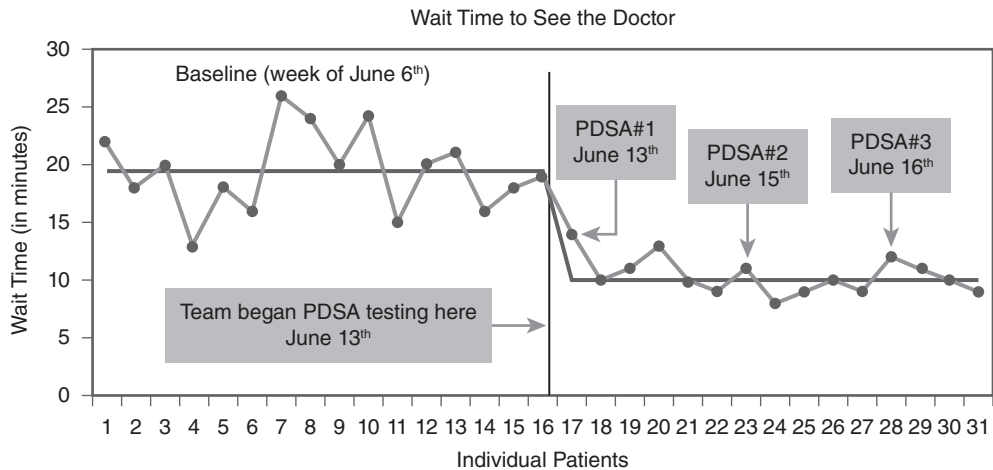
A challenge with any form of data analysis in this day and age is that it is very easy to push a button on a computer and get a variety of numbers, charts, and graphs. Some of these displays of data are very intricate and colorful. Some are three dimensional and some you can even rotate. Yet, I think we frequently rely on the machine at our fingertips and not the one that rests upon our shoulders. Pushing the buttons on the computer to get statistical and graphical results is in many ways the easy answer. Computers produce data not information. The goal is to turn data into information (Austin, 1983), which requires the machine on the top of your shoulders to be functioning fully and engaged.

Annotating your charts is a perfect way to start moving from data to information. Items that should be annotated on a chart include:

- The baseline period.
- The point at which the improvement team began testing new ideas. For example, annotating every Plan–Do–Study–Act (PDSA) cycle the team ran provides a well-documented journey.
- Annotating the detection of a run chart rule (e.g., a shift or trend) on the chart is very beneficial especially when presenting your work to a quality council or in a management meeting where there might be individuals who are not familiar with run chart analysis and interpretation. Moments like this provide a good opportunity for you to help educate others in your organization about run charts. But without annotations the key learning points might be missed.
- Besides annotating the actual run chart other useful annotations include:
 - Making sure you have a full and complete title on the chart. This should include a primary title identifying the facility or unit where the improvement work is occurring. This is usually in a little larger font (e.g., a size 14 or 16).
 - The secondary title for a chart should indicate the indicator by name and the dates for the data (e.g., Percentage of Compliance with the Ventilator Bundle, January 2017–March 2018).
 - Try to leave space to insert the operational definition for the data. If the chart takes up most of the page of your handout you can always place the operational definition and data collection plan on the reverse side of the page. If you are using slides, try reducing the size of the chart a little so you can place a small footnote (in a text box) at the bottom of the slide summarizing the operational definition and data collection plan.
 - If the indicator is a percentage or a rate it is always a good idea to at least provide a note (usually at the bottom of the chart under the horizontal axis) that reports the minimum and maximum size of the denominators. You do not necessarily have to include a full data table, which often is entirely too small and hard to read, but indicating the size of the denominators (min to max) will usually be enough to keep people from asking about them.
 - Finally, I find it useful to put the name of the individual who made the chart and contact information (in about a size 10 font) somewhere on the chart in case there are follow-up questions about the chart or its interpretation. This also comes in handy when materials handed out at a meeting are sent around to people who did not attend the meeting and they would like to talk to someone about the chart.

FIGURE 8-13 provides an example of applying these annotation suggestions to a run chart.

Stay Healthy Family Practice Clinic
Wait Time to See the Doctor, June 6th through June 17th



Operational Definition: Wait Time to See the Doctor is defined as the amount of time (measured in whole minutes) from when the patient checks in at the reception desk until he or she is seen by the doctor in the exam room.

Data Collection Plan: A systematic sample of 5 patients is pulled on 3 randomly selected days each week. Data are pulled during regular office hours. Weekends and scheduled holidays are excluded.

This chart was prepared by R. Lloyd who can be reached at Extension 1234.

FIGURE 8-13 Example of an annotated run chart

Notes

1. There is frequent confusion over the difference between the mean and the median. They are both measures of central tendency (i.e., they identify where a distribution of data is centered), but they are different. The mean is the arithmetic average (you add up all the numbers and divide by the total number of data points). The median, on the other hand, is the midpoint of a distribution of data points. This is the numerical value that divides the dataset exactly in half, the 50th percentile. The run chart uses the median as a centerline, but the control chart uses the mean as the centerline.
2. The data points on a run or Shewhart chart are usually not connected when you are interested in comparing the variation among members of a group (e.g., individual hospitals, clinics, or doctors) at a fixed point in time. For example, imagine that you are part of a nine-hospital system and want to see how all nine hospitals compare on the percentage of patients being properly assessed for the risk of a fall. Are all our hospitals performing as a system or do we have some doing very well and others not so well? In this case, you would collect data on each of the nine hospitals for 15 to 20 months and then make a run chart where each hospital

would be arranged along the horizontal axis of the run chart. The hospitals could be in alphabetical order, for example, or by bed size. The order of the units being compared would not matter for the purposes of comparison if the indicator is a percentage. Each hospital's data would be plotted and because this is not data over time but in the aggregate we would not connect the dots with a line. Typically, connecting the dots with a line indicates that this is a process that is flowing over time. The objective is to understand the variation amongst the dots irrespective of whether you connect them or not. An example of not connecting the dots is shown in Chapter 10, Case Study 1.

3. Although it is standard practice to place the median on a run chart as centerline (Carey and Lloyd, 2001; Provost and Murray, 2011; Pyzdek, 1990), some writers (e.g., Torki, 1992) advocate the use of the mean as the centerline on a run chart. The median is preferred by most writers because it is a measure of central tendency (along with the mean and mode) but more important it is not sensitive to extreme values as is the mean. All that matters with using the median is whether a data point is on this side of the median or the other side. If a data point is a small amount above the median or a very large distance above the median the median value will not be affected. This is because the median is positional (i.e., the 50th percentile point where half the data are above this point and half the data are below it). For the same reason, the median is not affected by data that are not normally distributed (i.e., suffering from skewness or kurtosis).
4. I frequently get asked in class, "Why do we add 1 to the number of crossings?" The answer is simple. Think of cutting a loaf of bread. If you cut the loaf of bread 10 times you actually end up with 11 slices of bread. The first cut actually produces two pieces of bread; one is the end heel cut and the rest of the loaf is actually the second slice. Similarly when the data line crosses the median the first time you have two segments of day. Adding 1 to the final count of the number of crossing accounts for the two segments that were created with the first crossing. You can verify this the next time you cut a loaf of bread. Count the number of times you have sliced the loaf and you will discover that you have one more slice of bread than the number of cuts you made.
5. I think a lot of people who don't have formal training in statistics think that it is a rather clear and definitive discipline. It uses numbers, formulae, Greek letters, and a variety of symbols that certainly make statistics and math look precise. We even refer to the precision of the p-value and talk about the .05 or even .01 level of significance. But there is a lot of gray area and debate in the field of statistics. A classic reference on this topic appeared in 1970. It is called *The Test of Significance Controversy* edited by Denton Morrison and Ramon Henkel (1970). If you are a student of statistical thinking this collection of readings is a must. A more contemporary and critical exploration of this topic is *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice and Lives* by Stephen Ziliak and Deirdre McCloskey (2011). I have been fascinated by this debate for years. My family knows this and at Christmas they bought me a sweatshirt that captures the ongoing debate nicely. It states in simple letters on the front of the sweatshirt "Statistics mean never having to be certain."
6. When you have a high and a low data point on a chart or even a couple high and several low points they essentially form the tails of a distribution. Think of the normally distributed bell-shaped curve. As you go out the tails of the normal distribution you expect to find less and less data, but

you will find a high and low data point residing in the tails of the distribution. But if you had only one extreme data point in a distribution it would not take the shape of the normal bell curve. In this case, you would have a skewed distribution with one rather long tail. The astronomical data point rule is similar to finding a nonnormal curve with one very long tail.

References

- Austin, C. *Information Systems for Hospital Administration*. Chicago: Health Administration Press, 1983.
- Lloyd, R. "Navigating in the Turbulent Sea of Data: The Quality Measurement Journey." Special edition on Quality Improvement in Neonatology and Perinatal Medicine, *Clinics in Perinatology* 37, no. 1 (March 2010): 101–122.
- Morrison, D., and R. Henkel, eds. *The Test of Significance Controversy*. Chicago: Aldine, 1970.
- Provost, L., and S. Murray. *The Health Care Data Guide*. San Francisco: Jossey-Bass, 2011.
- Pyzdek, T. *Pyzdek's Guide to SPC*. Vol. 1: *Fundamentals*. Milwaukee: Quality Press, 1990.
- Reason, J. *Human Error*. Cambridge: Cambridge University Press, 2003.
- Schultz, L. *Profiles in Quality*. New York: Quality Resources, 1994.
- Swed, F., and C. Eisenhart. "Tables for Testing Randomness of Grouping in a Sequence of Alternatives." *Annals of Mathematical Statistics* 14 (1943): 66–87.
- Western Electric Company. *Statistical Quality Control Handbook*. Indianapolis: AT&T Technologies, Inc., 1985.
- Wheeler, D., and D. Chambers. *Understanding Statistical Process Control*. Knoxville, TN: SPC Press, 1992.
- Ziliak, S., and D. McCloskey. *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice and Lives*. Ann Arbor: University of Michigan Press, 2011.

