Control Charts 101: A Guide to Health Care Applications

Sandra G. Amin

A control chart can be an effective way to display process data over time. It can differentiate common cause variation from special cause variation through the application of several probability-based interpretation rules. In addition, a control chart can be helpful in evaluating the effectiveness of a change. There are numerous types of control charts, and this tutorial was developed to guide the user through the selection process.

Key words: *control chart, data display, quality measurement, statistical process control*

Introduction

Purpose of Tutorial

This guide was developed in an attempt to provide an overview of control chart applications for common health care data. It was formulated with the presumption that the user has a basic understanding of process variation (i.e., is aware that variation is inherent in all processes and understands the difference between common cause variation and special cause variation), has knowledge of simple statistics (i.e., measures of central tendency), and grasps why one might want to look at a process using a control chart. When the decision has been made to use a control chart, this tutorial should help the user select the appropriate type of chart and understand the common rules of interpretation.

What is a control chart?

A control chart is a graphical display of data over time that can differentiate common cause variation (i.e., the variation that is inherent in every process and is due to chance) from special cause variation (i.e., variation in a process that is due to some assignable cause outside the process). In the late 1920s, Walter Shewhart, a statistician at the AT&T Bell Laboratories, developed the control chart and its associated rules of interpretation. These rules identify

***Sandra G. Amin, RN, MBA, CPHQ, CIC,** is Director, Quality Management, MetroHealth Medical Center, Cleveland, Ohio.*

This article is based on a workbook that was developed by the author and is used to teach staff at the MetroHealth Medical Center about the use of control charts. The workbook has been modified for publication, and comments should be directed to the author.

Quality Management in Health Care, 2001, 9(3), 1–27

special cause variation and are based on the probability that the graphical pattern would be very unlikely to occur by chance alone.

Control charts consist of a run chart displaying the measured data over time, a center line ($\overline{X}$), an upper control limit (UCL), and a lower control limit (LCL) (Figure 1). Upper and lower control limits are developed using mathematical formulas that are specific to the type of control chart that is selected to display the data. In general, they establish the margins within which the measurement will be found approximately 99% of the time.

A control chart can be an effective way to look at a process over time and to identify the presence of special cause variation. This useful tool can help identify the appropriate actions to take when wanting to improve a process (i.e., if special cause variation exists, it is important to remove it before beginning work to improve the process itself). In addition, a control chart can be useful when evaluating the effectiveness of change.

As noted earlier, the control limits are values calculated according to the type of data being graphed. The formulas are specific to the particular control chart chosen. Sometimes the LCL turns out to be a result that is not feasible and should be set to zero. For instance, it is not possible to have a negative quantity of defects so the LCL should be set to zero instead of the calculated value. Likewise, sometimes the UCL may calculate a value that is higher than is feasible (e.g., greater than 100% or a higher number of defective units than are being sampled). When this occurs, the UCL should be set to the highest feasible number instead of the calculated value.

Guide to Using Control Charts

Control Chart Interpretation Rules

Control charts identify the presence of special cause variation through the use of several probability-based interpretation rules. This is accomplished by looking at the graphical display for the presence of particular patterns. If one or more of the patterns is seen, it indicates the presence of special cause variation. Some of the more common control chart interpretation rules apply to the entire chart and include the following:

- one or more data points above a UCL or below an LCL (Figure 2); this would be due to chance occurrence only 1 in 200 times or 0.5% of the time
- eight or more points in a row that fall on either side of the center line (also known as a run) (Figure 3); this is the equivalent of having a coin toss come up heads eight times in a row and the likelihood of this being due to chance is 1 in 256 or $p < 0.005$
- seven or more consecutive points (if there are 21 or more data points) or six or more consecutive points (if there are less than 21 data points) steadily increasing or decreasing (also know as a trend); two successive points of the same value do not break the rule, however, these successive identical values would be counted as only one of the six or seven points for the rule (Figure 4)

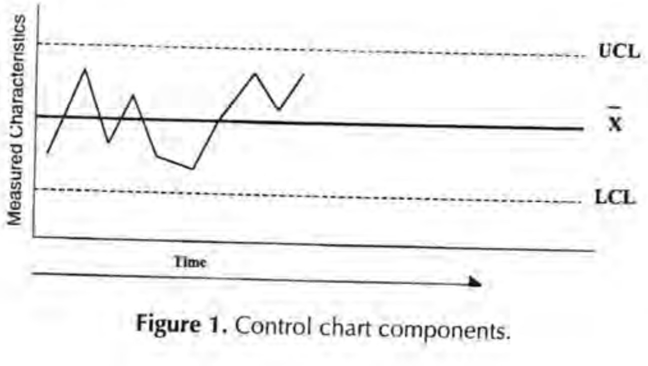


Figure 1. Control chart components.

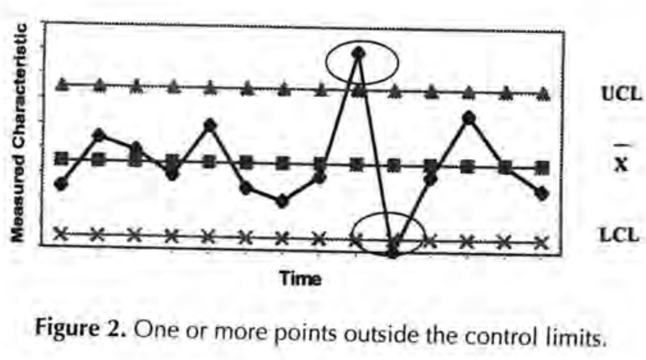


Figure 2. One or more points outside the control limits.

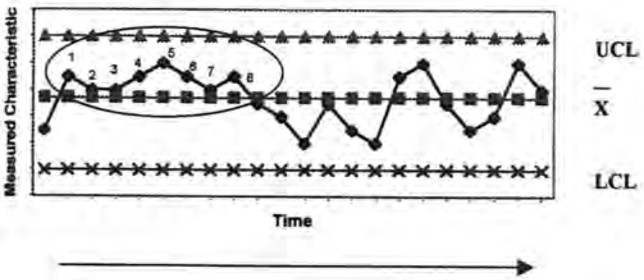


Figure 3. Eight or more points on the same side of the center line.

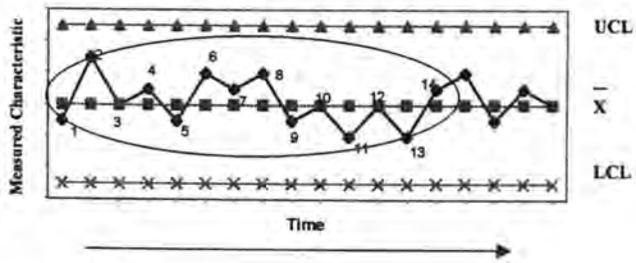


Figure 5. Fourteen or more points in a saw-tooth pattern.

- 14 or more consecutive points alternating up and down in a saw-tooth pattern (Figure 5)

For additional rules, the control chart is divided into six equal zones that fall between the UCL and LCL (Figure 6). Some of the additional rules that utilize these zones include:

- two of three successive points on the same side of the center line in Zone A
- four of five successive points on the same side of the center line in Zone B or beyond
- eight or more consecutive points occurring on either side of the center line and none of these points fall within Zone C (i.e., all fall in Zone B or beyond)
- 15 or more consecutive points occurring on either side of the center line that all fall within Zone C

Suggested Number of Data Points

It generally is recommended that there be at least 25 data points before constructing a control chart. However, if there are fewer than 25 points, it may be possible to:

- increase the number of data points by using a shorter time interval (i.e., use monthly or weekly data instead of quarterly data)
- begin the data display using a run chart instead of a control chart (refer to the section on this topic titled "Using a Run Chart to Display Data")
- use a control chart recognizing that control limits may be wider or narrower than is appropriate for the actual process due to the limited number of data points that were used in the calculations (e.g., the initial data points may not be fully representative of the process, etc.)

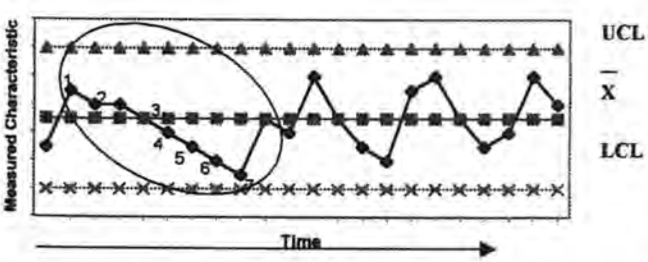


Figure 4. Seven or more successive points all decreasing.

Freezing and Revising Control Limits

Typically, the upper and lower control limits and the center line are recalculated with each successive data point that is added to the control chart. However, in addition to identifying special cause variation, a control chart can be a useful vehicle to evaluate the effectiveness of a change. Thus, when determining the impact of a process change, it may be helpful to maintain the control limits and center line from the old process and continue to plot the new process data points on the control chart. In this case, the identification of special cause variation indicates that the new process is significantly different from the old process (Figure 7). When there are enough data points from the new process to create a control chart, the control limits and center line are recalculated using only these new process data points and displayed as shown in Figure 8.

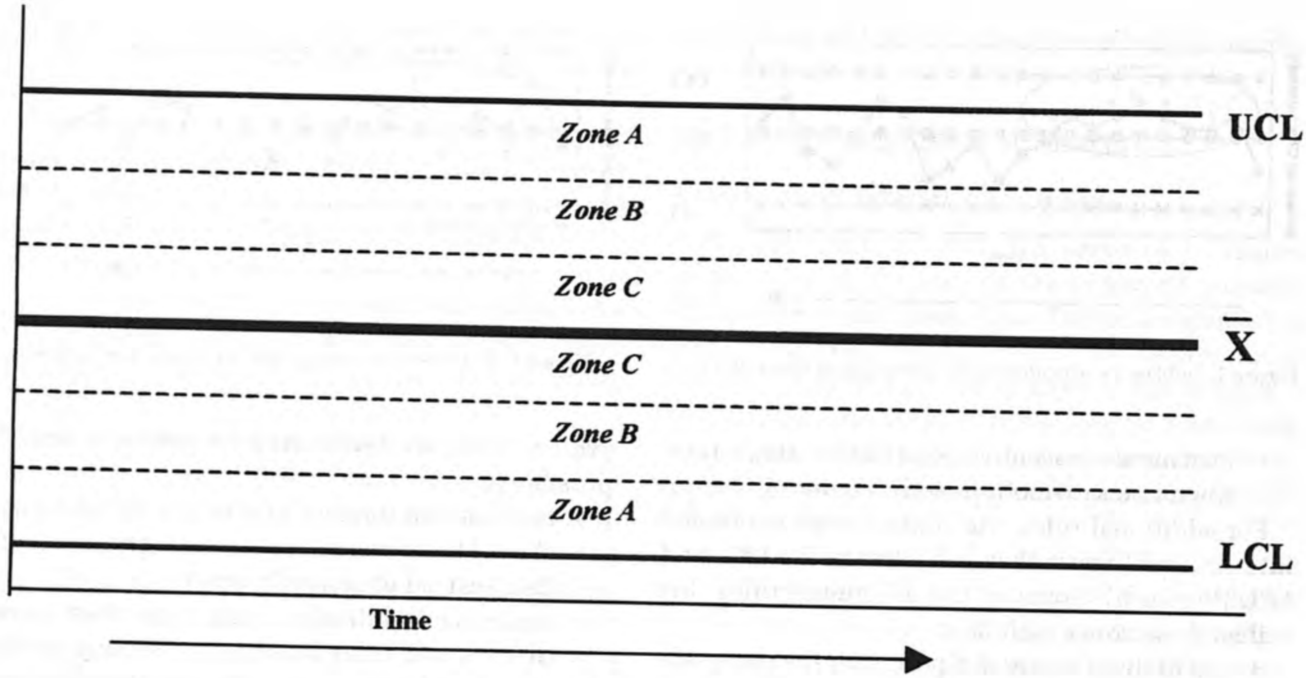


Figure 6. Control chart zones.

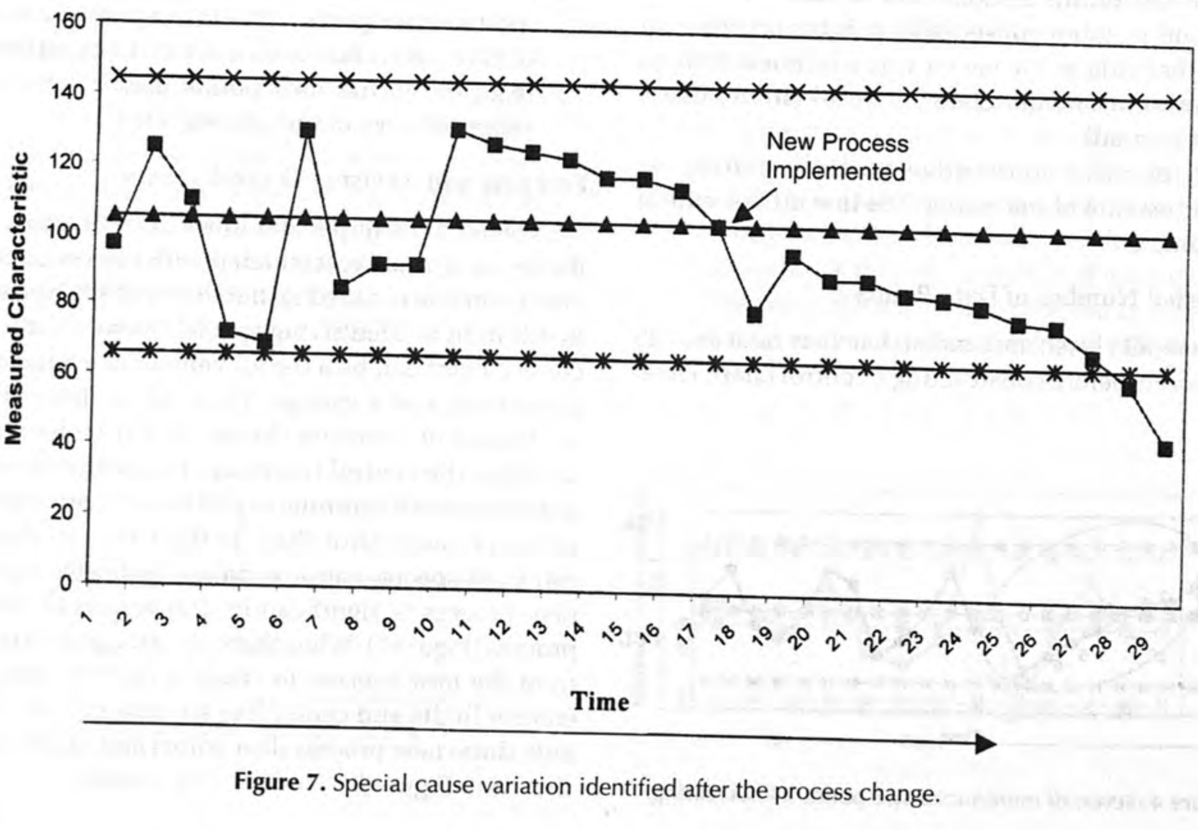


Figure 7. Special cause variation identified after the process change.

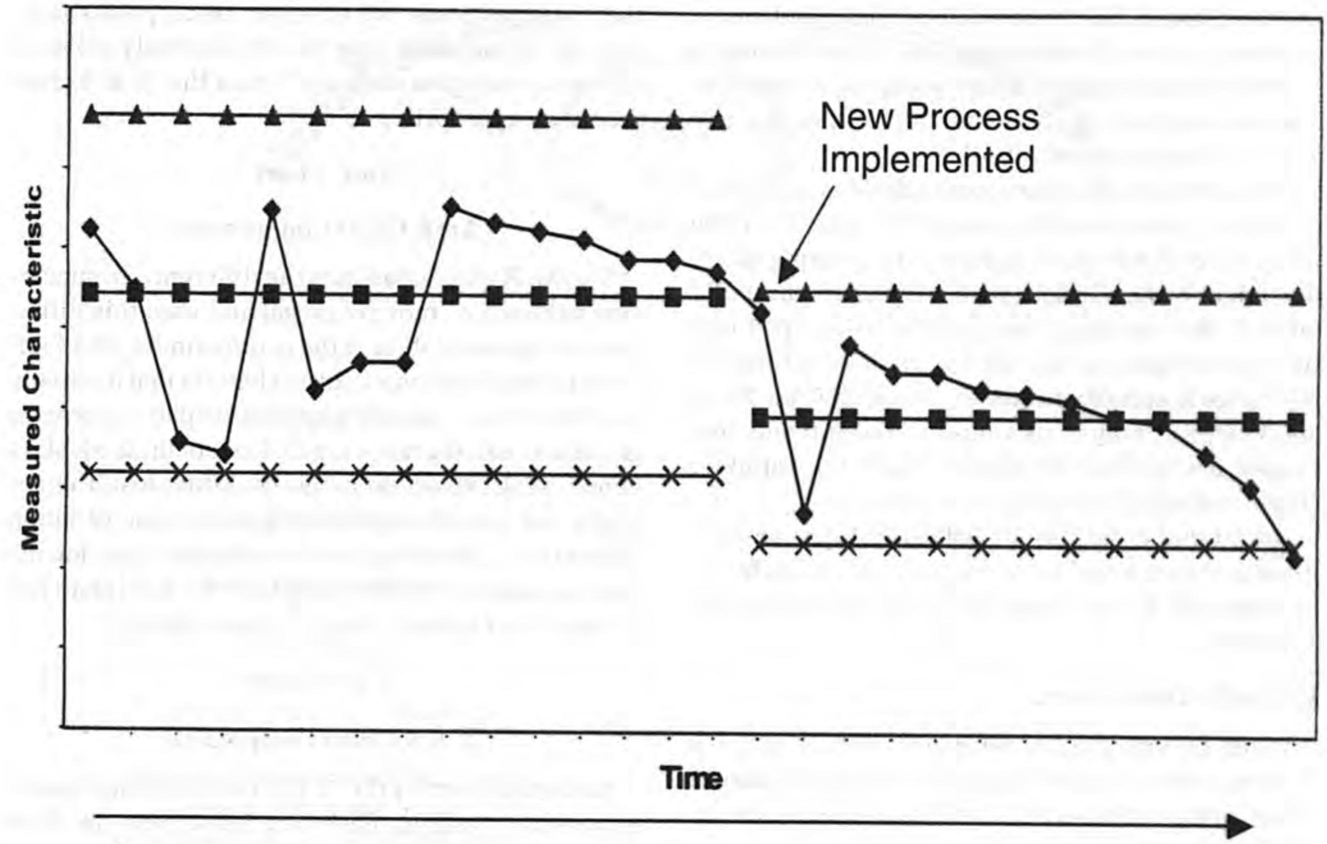


Figure 8. Control limits re-calculated for the new process.

Selecting the Appropriate Control Chart

While there are numerous types of control charts, seven charts are seen most commonly. The decision about which chart to use primarily rests with the type of data that have been collected and the methods used. This section of the tutorial is a guide through the process of selecting the correct control chart.

Types of Data

In order to be able to select the appropriate control chart to use, one first must be able to identify the type of data that are to be placed in the control chart. Data can be classified as either variables data or attributes data. There are three control charts that can be used for variables data (XmR chart, $\overline{X}$ & R chart, and $\overline{X}$ & S chart) and four control charts that can be used for attributes data (C chart, U chart, P chart, and NP chart). The thought process for determining which chart is appropriate will be covered in the Variables Chart and Attributes Chart Sections. Before getting to this, however, one needs a clear understanding of the type of data that one intends to display in a chart.

Variables data sometimes also are referred to as continuous data. These types of data generally represent *measurements* on a continuous scale that have an infinite number of possible values. Examples include the time to complete a process, systolic blood pressure measurements, pounds of infectious waste generated, and length of stay.

Attributes data sometimes also are referred to as discrete data. These types of data represent *counts* of events that can be grouped into discrete groups such as died versus did not die, infected versus not infected, had a fall versus did not fall, and had a cesarean section (C-section) versus did not have a C-section. These counts then can be expressed either as a whole number (e.g., number of deaths) or as a rate (e.g., % of patients who died).

Attributes data that have a considerable number of possible values resemble variables data from the perspective of statistical analysis. An example where this might be true is data such as monthly mortality rates. In this example, the data can be grouped into two discrete groups (i.e., died or did not die) and the rate range is actually finite (i.e., 0% to 100%). There are, however, numerous values between 0 and 100. Consequently, it is acceptable to use the variables charts instead of the attributes charts.

Additional examples for both data types are outlined in Figure 9 and a practice exercise can be found in Appendix 1. See Appendix 4 for Practice Review Answers.

Variables Data Charts

When a researcher has determined that he or she is dealing with variables data, the next step is to decide which of the variables data control charts is appropriate. There are three possible choices:

- XmR (individuals and moving range chart)
- $\overline{X}$ & R (average & range chart or X bar & range chart)
- $\overline{X}$ & S (average & sigma chart or X bar & sigma chart)

The appropriate chart is selected by examining one of the process data points (Figure 10). If that point is a single measure (e.g., turnaround time for a single laboratory test), then the XmR chart is appropriate. If that point is made up of a set of data that contains more than 1 point and less than or equal to 10 points (e.g., the average turnaround time for 6 laboratory tests), then the $\overline{X}$ & R chart would be appropriate. If that point is made up of 11 or more points (e.g., average turnaround time for 20 randomly selected laboratory samples each day), then the $\overline{X}$ & S chart would be appropriate.

> ***The appropriate chart is selected by examining one of the process data points.***

XmR Chart

XmR Chart Components

The XmR chart considers the difference in successive values (i.e., moving range) and uses this difference in the calculation of the control limits. The XmR chart differs from some control charts in that it actually consists of two separate graphical displays, a process graph and moving range graph. Each of these displays contains calculated control limits and as a result, either graph can provide signals for special cause variation (Figure 11). Formulas and construction tips for the moving range and process graphs of the XmR chart can be found in Figures 12 and 13, respectively.

$\overline{X}$ & R Chart

$\overline{X}$ & R Chart Components

Each data point of the $\overline{X}$ & R chart is comprised of groups of subvalues. Each data point consists of the same number of subvalues and an $\overline{X}$ & R chart is the appropriate chart when there are at least two subvalues but not more than 10 subvalues per data point. The chart is created by considering each group's average and range values. The control limits then are calculated using these values and a table of constants (Figure 14) that correspond to the number of subvalues per data point. Like the XmR chart, the $\overline{X}$ & R chart consists of two (2) separate graphical displays, a process graph and a range graph. Each of these displays contains calculated control limits, and as a result either graph can provide signals for special cause variation (Figure 15). Formulas and construction tips for the range and process portions of the $\overline{X}$ & R chart can be found in Figures 16 and 17, respectively.

$\overline{X}$ & S Chart

$\overline{X}$ & S Chart Components

Each data point of the $\overline{X}$ & S chart consists of 11 or more subvalues. Like the $\overline{X}$ & R chart, each data point

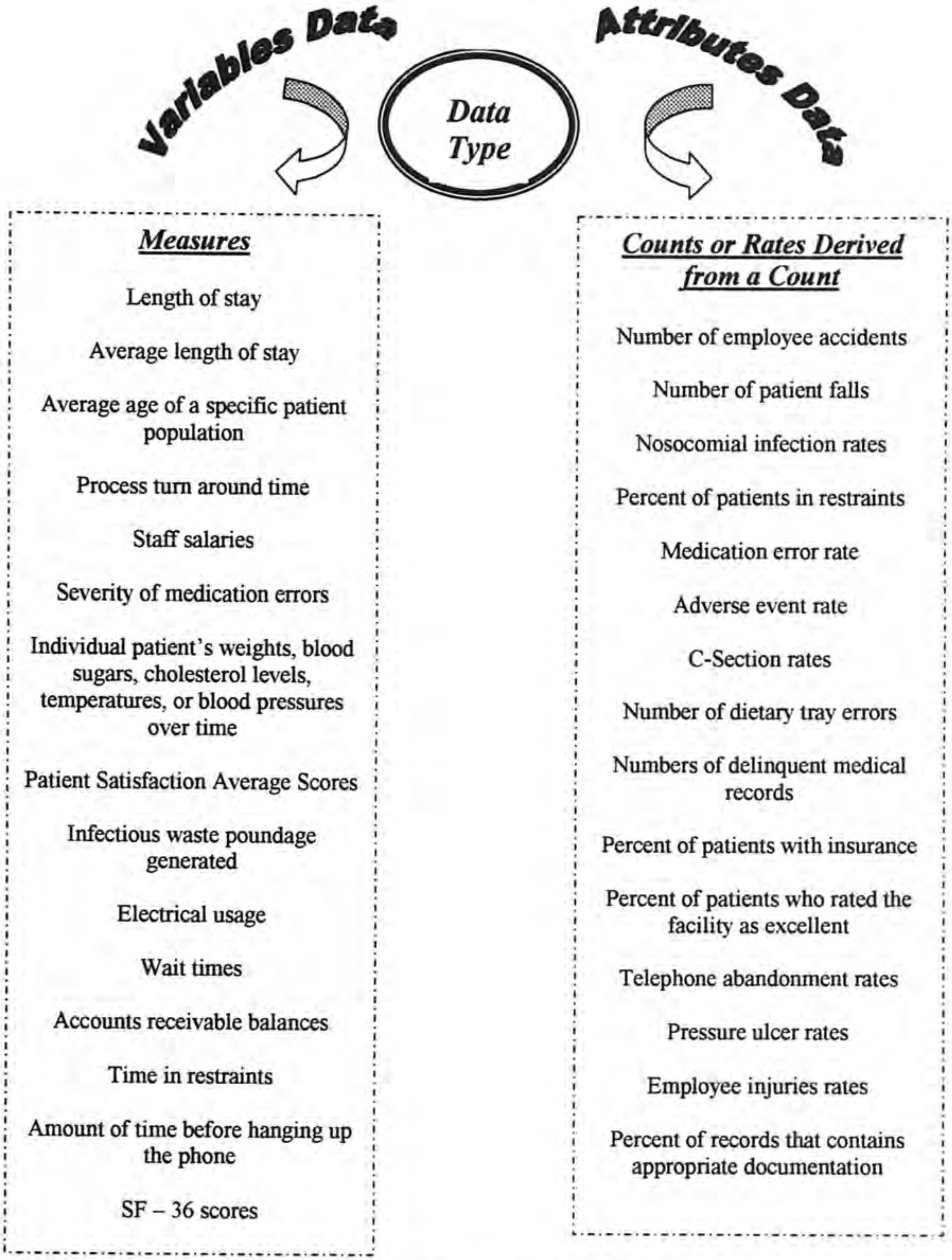


Figure 9. Examples of variables and attributes data.

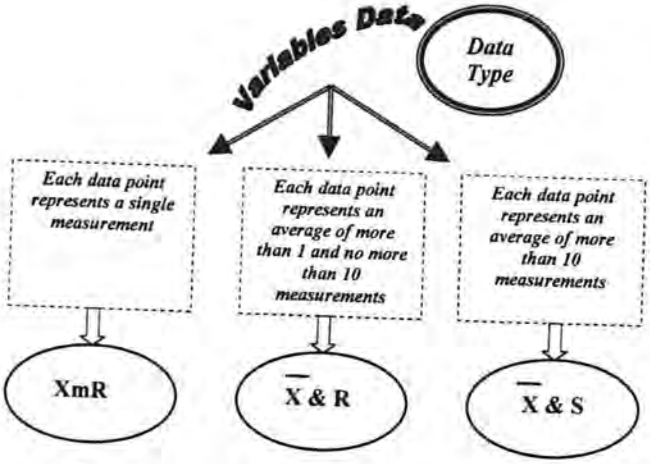


Figure 10. Types of variables data control charts.

is comprised of the same number of subvalues. The control limits are calculated using each subgroup's average value and standard deviation. Like the $\overline{X}$ & R chart, the $\overline{X}$ & S chart has two graphical displays (i.e., a process graph and a sigma graph) (Figure 18) and uses a table of constants (Figure 19) for the calculations of the control limits. Formulas and construction tips for the sigma and process portions of the $\overline{X}$ & S chart are covered in Figures 20 and 21, respectively.

Attributes Data Charts

When researchers have determined that they are dealing with attributes data, the next step is to decide which of the four attributes charts is most appropriate. The possible choices are:

- C chart
- U chart
- P chart
- NP chart

The appropriate chart is selected by first determining if the data points are counts of the number of defects/number of non-conformances or are counts of defectives/non-conforming units (Figure 22). Sometimes it is helpful to ask if the event could happen more than once to the same item. If the answer is "yes" (e.g., falls, injuries, number of errors on a dietary tray, etc.), then researchers are counting the number of defects or number of non-conformances. If the answer is "no" (e.g., deaths, C-section delivery for this pregnancy, etc.), then they are counting defectives or non-conforming units. Additionally, if researchers are comparing an item to a standard and then classifying it as whether it meets the standard or not, then they are counting defectives or non-conforming units (e.g., counting the number of dietary trays that are 100% correct, etc.).

If the data are derived from counts of defects or non-conformances, then the next step is to determine whether to use the C chart or the U chart (Figure 23). If the sample size or the population at risk is the same for each data point (e.g., each data point represents the number of errors in a sample of 25 dietary trays), then a C chart is used. If the sample size or population at risk varies significantly for each data point (e.g., events per monthly discharges, etc.), then a U chart is generally appropriate. Sometimes if the sample size varies only slightly from time period to time period (e.g., the number of discharges is relatively stable), then a C chart could be used instead of a U chart.

If the data points are derived from counts of defectives or non-conforming items, then the next step is to determine whether to use the P chart or the NP chart (Figure 24). If the sample size or population at risk varies for each data point, then the P chart should be used (e.g., monthly C-section rates because the total number of deliveries would vary from month to month). If the sample size or population at risk for each time period is the same, then an NP chart would be appropriate (e.g., number of perfect dietary trays out of a random sample of 50 trays per month).

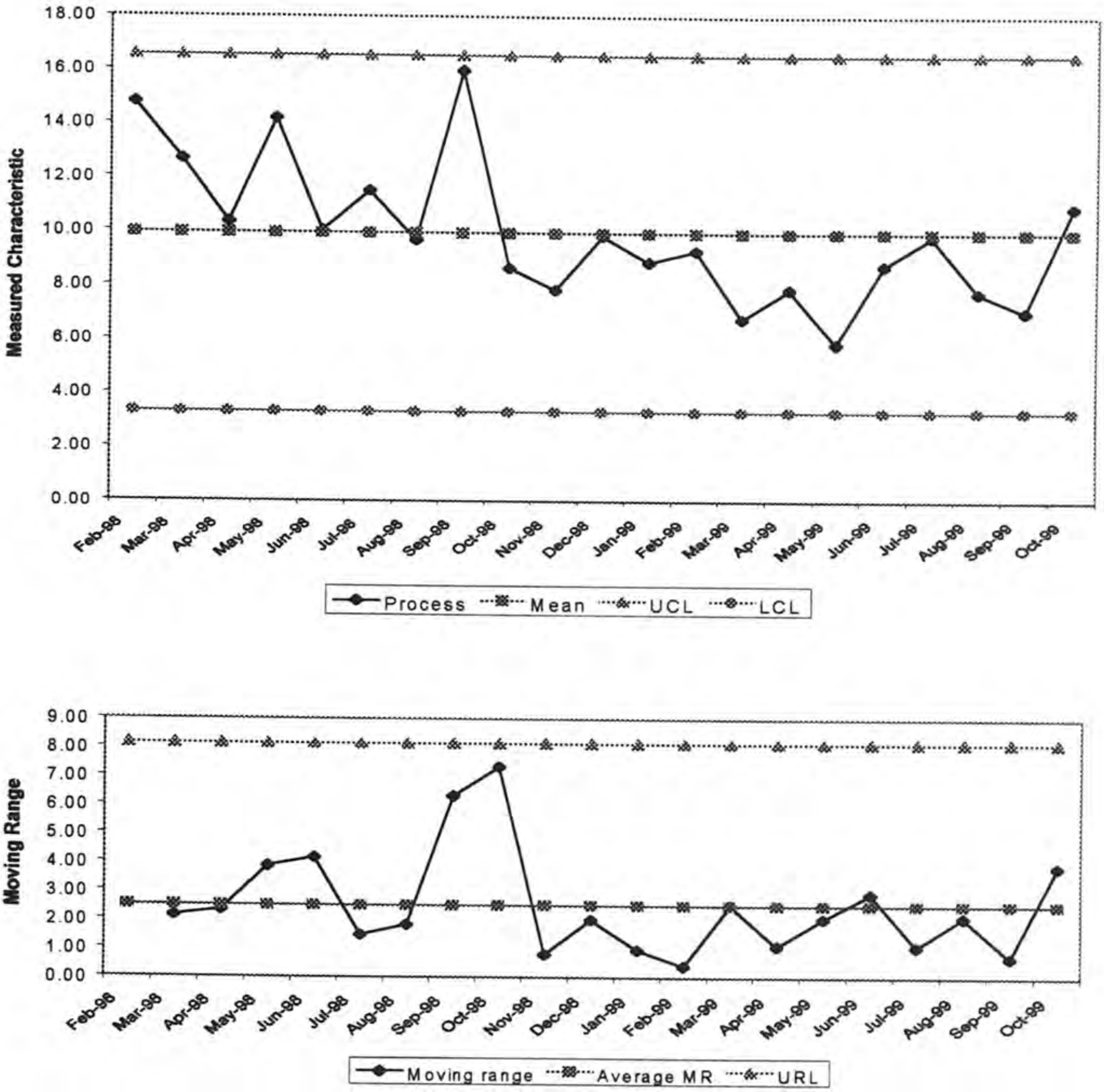


Figure 11. Sample XmR chart.

Moving Range Portion of XmR Control Chart

- Place the process values (**X**) in a time-ordered sequence
- Calculate the difference between successive values
- Take the absolute value of the resulting difference and call this value the moving range (**mR**). **Note**: there will be one fewer mR values than there are process values.
- Calculate the average of all the moving ranges ($\overline{\mathbf{mR}}$)

$$\overline{\mathbf{mR}} = \frac{\text{sum of the mR values}}{\text{\# of mR values}}$$

- Calculate the upper range limit (**URL**)

$$\mathbf{URL} = 3.27 \times \overline{\mathbf{mR}}$$

- Plot the **mR**, $\overline{\mathbf{mR}}$ and **URL** values on the same graph, and place the graph below the process graph.

Figure 12. Calculations for the moving range portion of XmR chart.

Process Portion of XmR Control Chart

- Calculate the average of the process values ($\overline{\mathbf{X}}$)

$$\overline{\mathbf{X}} = \frac{\text{Sum of the process values}}{\text{\# of process values}}$$

- Calculate the upper control limit (**UCL**) **Note:** If the UCL calculation yields a value that is higher than is possible, then set the UCL at the maximum possible value.

$$\mathbf{UCL} = \overline{\mathbf{X}} + (2.66 \times \overline{\mathbf{mR}})$$

- Calculate the lower control limit (**LCL**) **Note**: If the LCL < 0, then set the LCL = 0.

$$\mathbf{LCL} = \overline{\mathbf{X}} - (2.66 \times \overline{\mathbf{mR}})$$

- Plot the **X**, $\overline{\mathbf{X}}$, **UCL**, and **LCL** values on the same graph, and place the graph above the range graph.

Figure 13. Calculations for the process portion of XmR chart.

Appendix 2 contains a practice exercise for selecting the appropriate control chart for both variables data and attributes data. See Appendix 5 for Practice Review Answers.

C Chart

C Chart Components

The C chart considers the process average and calculates its control limits from this value. Both counts and rates can be displayed in a C chart and it is presumed that the number of defects is rare and that these data have a Poisson probability distribution (i.e., distribution of rare events). Unlike the variables charts, there is only one graph associated with the C chart (Figure 25). Formulas and construction tips for a C chart can be found in Figure 26.

U Chart

U Chart Components

The U chart considers the sample size for each data point and uses this sample size in the calculation of

Formulas for the range & process charts utilize coefficient factors that vary according to the size of the sub-groups that make up each of the data points on the chart. The coefficient factors are as follows:

Size of Sub-group	Factor for Process Chart (P)	Factor for Range Chart LCL (L)	Factor for Range Chart UCL (U)
2	1.88	0.00	3.27
3	1.02	0.00	2.57
4	0.73	0.00	2.28
5	0.58	0.00	2.11
6	0.48	0.00	2.00
7	0.42	0.08	1.92
8	0.37	0.14	1.86
9	0.34	0.18	1.82
10	0.31	0.22	1.78

Figure 14. Table of constants for $\overline{X}$ & R chart.

$\overline{X}$ & R Chart Formulas and Construction Tips

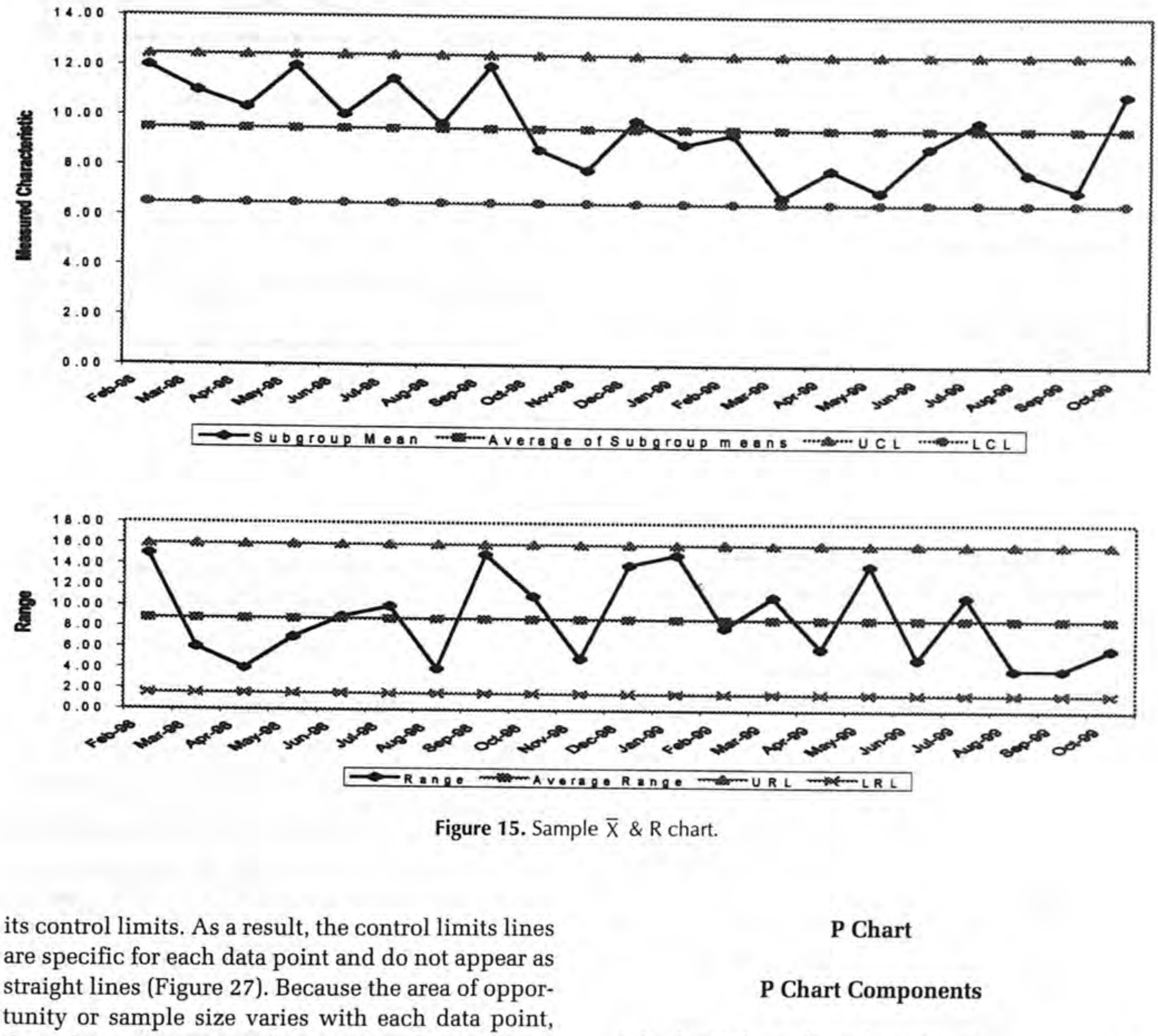


Figure 15. Sample $\overline{X}$ & R chart.

its control limits. As a result, the control limits lines are specific for each data point and do not appear as straight lines (Figure 27). Because the area of opportunity or sample size varies with each data point, these data points must reflect a rate and not a count. If a count were used, the resulting data would be meaningless because the count would not control for changes in the sample size. As with the C chart, there is only one graph and the data should follow a Poisson probability model (i.e., distribution of rare events). Formulas and construction tips for a U chart are shown in Figure 28.

P Chart

P Chart Components

Like the U chart, the P chart data elements must be a rate because the sample size or opportunity for the occurrence varies for each data point. The control limits take this variation in sample size into consideration and thus the control limits do not appear as straight lines (Figure 29). The P chart is most applicable when the rate of defectives is not rare (e.g., > 5%) and it is assumed that the data follow the

Range Portion (R) of the $\overline{X}$ & R Control Chart

- Calculate the range (R) for each sub-group

  R = (maximum value in sub-group) — (minimum value in sub-group)

- Calculate the average range for all data points

$$\overline{R} = \frac{R_1 + R_2 + \ldots R_n}{\text{number of data points}}$$

- Calculate the upper range limit (**URL**)

  URL = (factor in "U" column that corresponds to the sub-group size) **x** $\overline{R}$

- Calculate lower range limit (**LRL**)

  LRL = (factor in "L" column that corresponds to the sub-group size) **x** $\overline{R}$

- Plot **R, $\overline{R}$, URL & LRL** on the same graph, and place the graph below the process graph

Figure 16. Calculations for range portion of $\overline{X}$ & R chart.

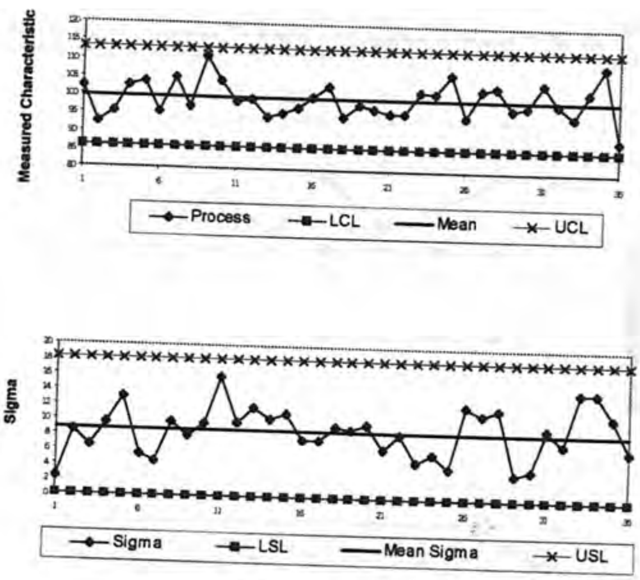


Figure 18. Sample $\overline{X}$ & S chart.

Process Portion $(\overline{X})$ of the $\overline{X}$ & R Control Chart

- Calculate the average value $(\overline{X})$ of the sub-values (X) for each data point

$$\overline{X} = \frac{X_1 + X_2 + \ldots X_n}{\text{Number of sub-values for the data points}}$$

- Calculate the average of the average values for the sub-values $(\overline{\overline{X}})$ for each data point

$$\overline{\overline{X}} = \frac{\overline{X_1} + \overline{X_2} + \ldots \overline{X_n}}{\text{Number of data points}}$$

- Calculate the upper control limit (UCL) <u>**Note:**</u> If the UCL calculation yields a value that is higher than is possible, then set the UCL at the maximum possible value.

  UCL = $\overline{\overline{X}}$ + [(factor in "P" column that corresponds to the sub-group size) **x** $\overline{R}$]

- Calculate the lower control limit (LCL) <u>**Note:**</u> If the LCL < 0, then set the LCL at zero.

  LCL = $\overline{\overline{X}}$ - [(factor in "P" column that corresponds to the sub-group size) **x** $\overline{R}$]

- Plot **$\overline{X}$, $\overline{\overline{X}}$, UCL & LCL** on the same graph, and place the graph above the range graph

Figure 17. Calculations for process portion of $\overline{X}$ & R chart.

binomial probability model. Formulas and construction tips for a P chart can be found in Figure 30.

NP Chart

NP Chart Components

The NP chart is used when looking at items that are compared with a standard and then are classified as to whether or not they meet the standard. The sample size (e.g., the number of units checked) is the same for each data point and as a result, the data point can be expressed as either a number or a rate (Figure 31). Like the P chart, NP chart data should follow the binomial probability model. Formulas and construction tips for an NP chart are outlined in Figure 32.

Can More Than One Chart Be Correct?

Admittedly, the preceding sections of this tutorial and the overview chart of the entire selection process (Appendix 3) make it seem as if the process of

Formulas for the sigma & process charts utilize coefficient factors that vary according to the size of the sub-groups that make up each of the data points on the chart. The coefficient factors are as follows:

Size of Sub-group	Factor for Process Chart (P)	Factor for Sigma Chart LCL (L)	Factor for Sigma Chart UCL (U)
11	0.93	0.32	1.68
12	0.89	0.35	1.65
13	0.85	0.38	1.62
14	0.82	0.41	1.59
15	0.79	0.43	1.57
16	0.76	0.45	1.55
17	0.74	0.47	1.53
18	0.72	0.48	1.52
19	0.70	0.50	1.50
20	0.68	0.51	1.49

Figure 19. Table of constants for $\overline{X}$ & S chart.

Sigma Portion (S) of the $\overline{X}$ & S Control Chart

- Calculate the average value $(\overline{X})$ for each data point sub-group.

$$\overline{X} = \frac{X_1 + X_2 + \ldots X_n}{\text{Number of sub-values}}$$

- Calculate the standard deviation (S) for each data point sub-group.

$$S = \sqrt{\frac{(X_1 - \overline{X})^2 + (X_2 - \overline{X})^2 + \ldots (X_n - \overline{X})^2}{(n-1)}}$$

- Calculate the average of the standard deviations $(\overline{S})$ for all of the sub-groups.

$$\overline{S} = \frac{S_1 + S_2 + \ldots S_n}{\text{Number of data points}}$$

- Calculate the upper sigma limit (USL)

USL = (factor in "U" column that corresponds to the sub-group size) x $\overline{S}$

- Calculate lower sigma limit (LSL)

LSL = (factor in "L" column that corresponds to the sub-group size) x $\overline{S}$

- Plot **S, $\overline{S}$, USL & LSL** on the same graph, and place the graph below the process graph

Figure 20. Calculations for sigma portion of $\overline{X}$ & S chart.

Process Portion $(\overline{X})$ of the $\overline{X}$ & S Control Chart

- Calculate the average value $(\overline{\overline{X}})$ for all of the sub-group averages $(\overline{X})$

$$\overline{\overline{X}} = \frac{\overline{X}_1 + \overline{X}_2 + \ldots \overline{X}_n}{\text{Number of data points}}$$

- Calculate the average of the standard deviations for the sub-groups $(\overline{S})$ for each data point

$$\overline{\overline{S}} = \frac{S_1 + S_2 + \ldots S_n}{\text{Number of data points}}$$

- Calculate the upper control limit (UCL) **Note:** If the UCL calculation yields a value that is higher than is possible, then set the UCL at the maximum possible value.

UCL = $\overline{\overline{X}}$ + [(factor in "P" column that corresponds to the sub-group size) **x** $\overline{S}$]

- Calculate the lower control limit (LCL) **Note:** If the LCL < 0, then set the LCL at zero.

LCL = $\overline{\overline{X}}$ - [(factor in "P" column that corresponds to the sub-group size) **x** $\overline{S}$]

- Plot **$\overline{X}$, $\overline{\overline{X}}$, UCL & LCL** on the same graph, and place the graph above the sigma graph

Figure 21. Calculations for process portion of $\overline{X}$ & S chart.

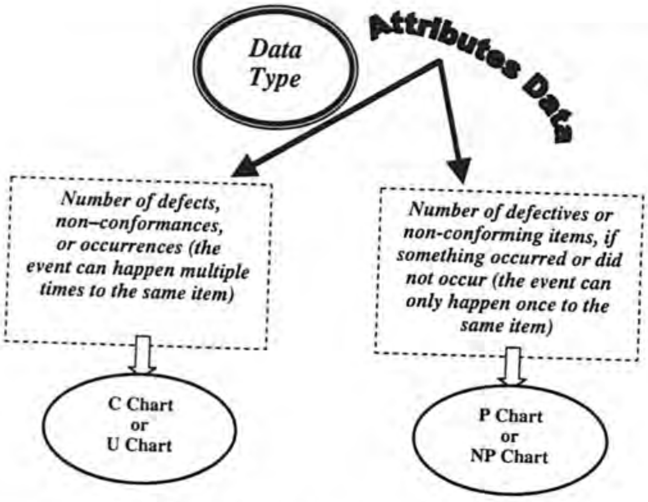


Figure 22. Types of attributes data control charts.

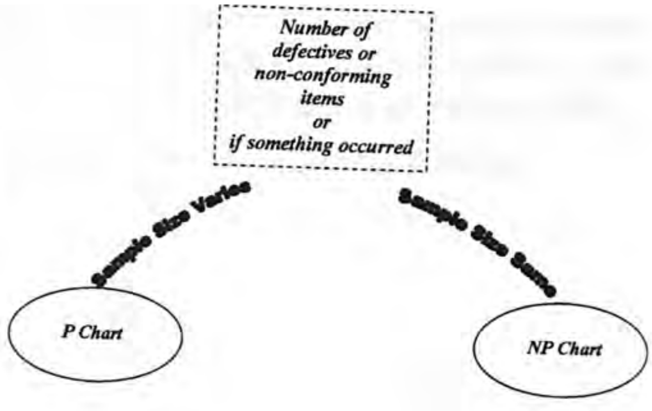


Figure 24. Control charts for defectives.

selecting the appropriate control chart is very black and white. In reality, sometimes more than one control chart type could be an appropriate selection. This typically occurs for a couple of reasons.

One cause of this phenomena is associated with the decision point where the type of data (i.e., variables versus attributes) is determined. The fact that it is permissible to use variables data charts when the attributes data have a large number of possible values expands the potential chart selection significantly.

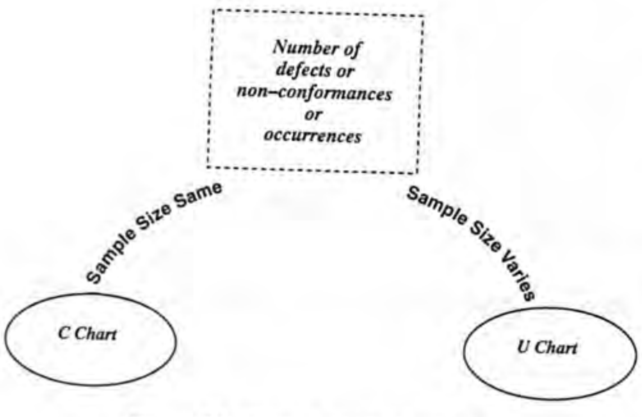


Figure 23. Control charts for defects.

Another consideration that can alter the chart selection process is the probability model that is associated with the data. The P and NP charts assume that the data have binomial probability and the C and U charts assume that the data have Poisson probability. Thus, if researchers do not know how to or care to evaluate the appropriateness of the binomial and/or Poisson probability models, then these charts probably should be avoided. The variables charts are not based on a specific probability model and thus can be used when the probability distribution is not known.

In addition, some judgment is used in determining if the sample sizes or populations at risk are the same for each time period. If the data were sampled in a specific manner (i.e., using a constant sample size), the choice is obvious. The variability comes into play when the population at risk is approximately the same from time period to time period (e.g., the number of discharges is approximately the same month to month). It then could be acceptable to consider that

In reality, sometimes more than one control chart type could be an appropriate selection.

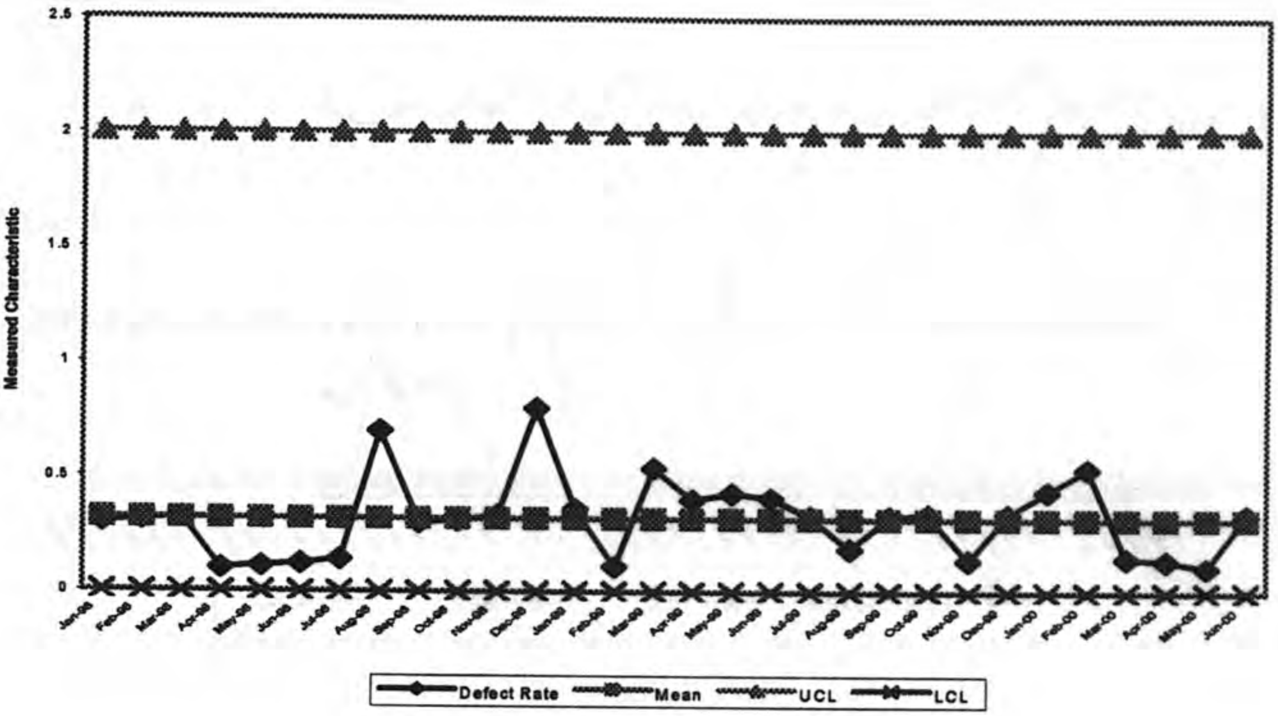


Figure 25. Sample C chart.

- If the data point is a count instead of a rate (i.e. C = # of defects), then calculate the average number of defects ($\overline{C}$) as

$$\overline{C} = \frac{\text{\# of defects}}{\text{\# of observations}}$$

- If the data point is a rate (i.e. C = # defects/# of observations), then calculate the average defect rate as

$$\overline{C} = \frac{\text{Sum of all defects}}{\text{Sum of all observations}}$$

- Calculate sigma (s)

$$s = \sqrt{\overline{C}}$$

- Calculate upper control limit (UCL)

$$\text{UCL} = \overline{C} + (3 \times s)$$

- Calculate lower control limit (LCL). Note: If the LCL < 0, then set the LCL to zero.

$$\text{LCL} = \overline{C} - (3 \times s)$$

- Plot C, $\overline{C}$, UCL, and LCL on the same graph

Figure 26. Calculations for C chart.

the sample size is constant from time period to time period.

Generally speaking, the above mentioned issues will not result in a failure to identify special cause variation when it in fact exists. An example whereby the same data were put into an XmR chart and a U chart can be found in Figures 33 and 34, respectively. Special cause variation was evident in both of these charts at the same point in time. Whichever chart is selected, it would be important to be able to explain how and defend why the chart was determined.

Using a Run Chart to Display Data

Run Chart Components

Like a control chart, a run chart illustrates data over time and can differentiate special cause variation from common cause variation. It looks similar to a control chart except that it does not include an upper or lower control limit (Figure 35). A run chart might be used when there are too few data points for a

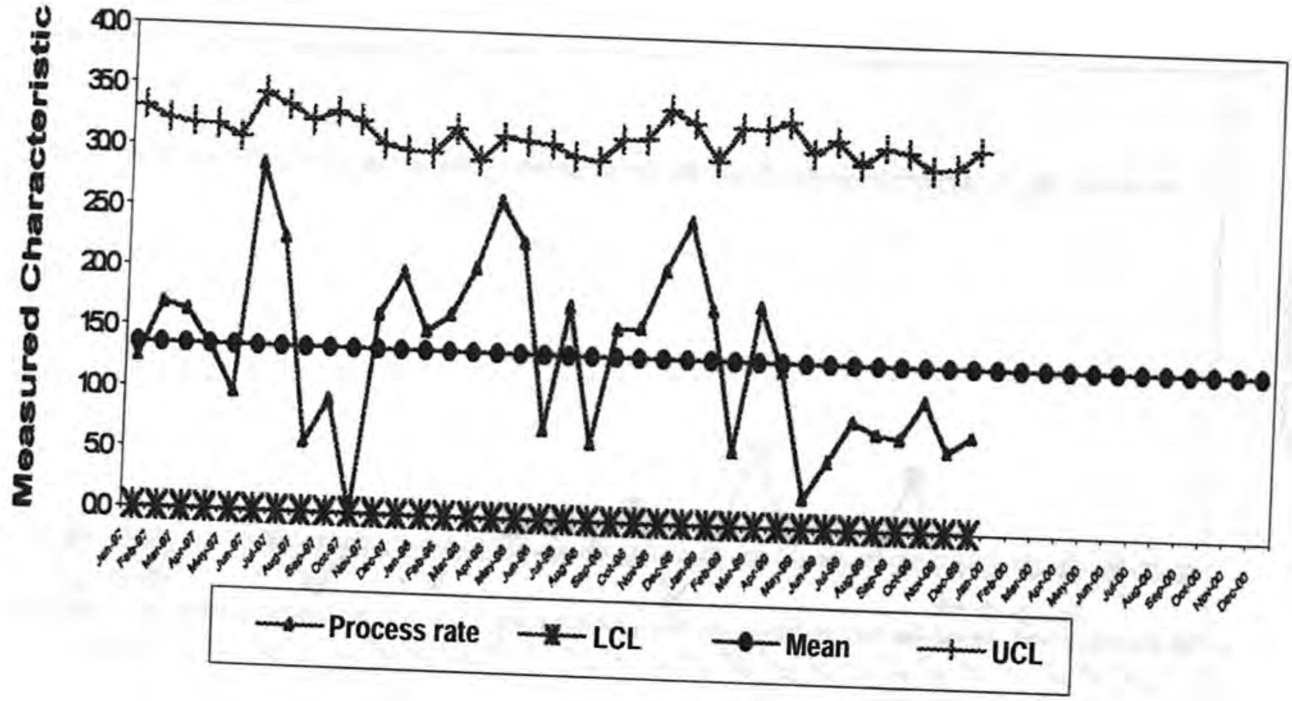


Figure 27. Sample U chart.

- Calculate each data point (U)

$$U = \frac{\text{\# of defects}}{\text{\# of observations}}$$

- Calculate the process average $(\overline{U})$

$$\overline{U} = \frac{\text{Total \# of defects for all data points}}{\text{Total \# of observations for all data points}}$$

- Calculate Sigma (*s*) for each data point. __Note__: *s* will vary as the # of observations for each data point changes

$$s = \sqrt{\frac{\overline{U}}{\text{Number of observations for each data point}}}$$

- Calculate the upper control limit (**UCL**) for each data point.

$$UCL = \overline{U} + (3 \times s)$$

- Calculate the lower control limit (**LCL**) for each data point. __Note__: If the calculated value **LCL** < **0**, then set the **LCL** to zero.

$$LCL = \overline{U} - (3 \times s)$$

- Plot **U**, $\overline{U}$, **UCL** and **LCL** on the same graph.

Figure 28. Calculations for U chart.

control chart. Additionally, it is quick and easy to construct and can identify special cause variation.

Run Chart Interpretation Rules

Run chart interpretation rules are similar to some of the control chart rules and, like control chart rules, they are based on the probability that certain patterns are unlikely to have occurred by chance alone. Indications of special cause variation are as follows:

- a series of points in a row (a run) that fall on either side of the center line: For this rule, the number of points needed in a row to indicate special cause varies according to the number of data points on the chart (i.e., less than 20 data points would require seven or more points in a row and 20 or more data points would require eight or more points in a row).
- consecutive points steadily increasing or decreasing (a trend): Two successive points of the same value do not break the rule, however, these successive identical values would be counted as only one of the points for the rule. The number of points needed to apply this rule also varies according to

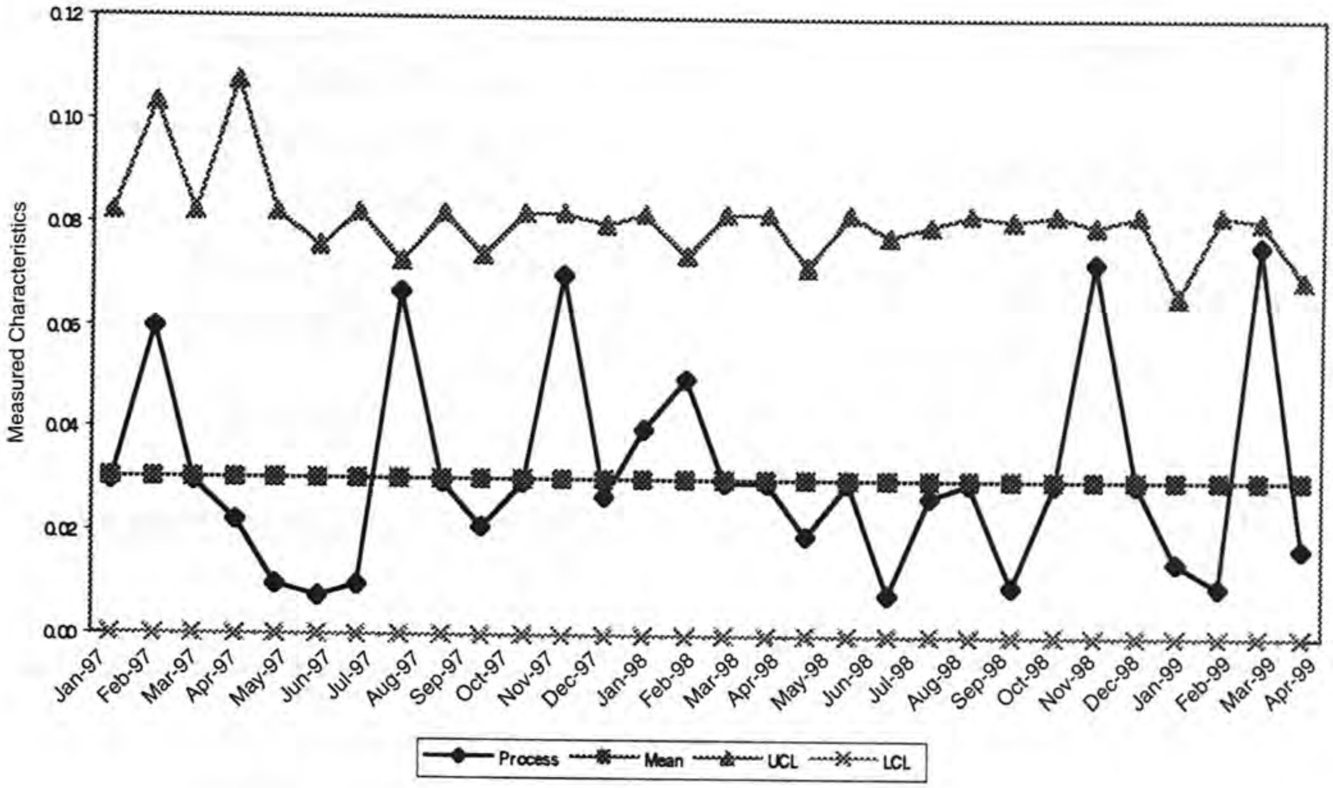


Figure 29. Sample P chart.

the number of data points on the chart (i.e., 5–8 data points would require 5 or more points in a row, 9–20 data points would require 6 or more points in a row, and 21–100 data points would require 7 or more points in a row).

- fourteen or more consecutive points alternating up and down in a saw-tooth pattern

Run Chart Formulas

Calculating the value for the center line ($\overline{X}$):

- If the process values represent a count of some parameter and do not represent a rate, the center line would be calculated by taking the average of the values (Figure 36).
- If the process values represent a rate (i.e., have components of both a numerator and denominator such as % mortality, # of events per 1000 patient days, etc.), the center line would be calculated by taking the sum of the numerators and dividing that value by the sum of the denominators (Figure 37).

Tools of the Trade

Control charts can be created easily using a basic spreadsheet software product such as Microsoft Excel. Additionally, numerous commercially available software packages have been developed to create the different types of control charts. There are advantages and disadvantages for each type of product and both need to be weighed when determining which product is most appropriate for your needs.

For example, spreadsheet programs are readily available on most computers. As a result, additional costs are avoided and files with the data and graphs can be shared easily with others. The ability to share

- Calculate each data point (**P**)

$$\mathbf{P} = \frac{\text{\# of defectives}}{\text{\# of occurrences}}$$

- Calculate the process average $(\overline{\mathbf{P}})$

$$\overline{\mathbf{P}} = \frac{\text{Total \# of defectives for all data points}}{\text{Total \# of occurrences for all data points}}$$

- Calculate Sigma (*s*) for each data point. __Note__: *s* will vary as the # of observations for each data point changes

$$s = \sqrt{\frac{\overline{\mathbf{P}} \times (1 - \overline{\mathbf{P}})}{\text{Denominator for the individual data point}}}$$

- Calculate the upper control limit (**UCL**) for each data point. __Note__: If the calculated **UCL** value is higher than is feasible, set the **UCL** to the highest feasible value.

$$\mathbf{UCL} = \overline{\mathbf{P}} + (3 \times s)$$

- Calculate the lower control limit (**LCL**) for each data point. __Note__: If the calculated value **LCL < 0**, then set the **LCL** to zero.

$$\mathbf{LCL} = \overline{\mathbf{P}} - (3 \times s)$$

- Plot **P**, $\overline{\mathbf{P}}$, **UCL** and **LCL** on the same graph.

Figure 30. Calculations for P chart.

- If the data point is a count instead of a rate (i.e. **nP** = # of defectives), then calculate the average number of defectives $(\mathbf{n\overline{P}})$ as

$$\mathbf{n\overline{P}} = \frac{\text{Total \# of defectives}}{\text{\# of data points}}$$

- If the data point is a rate (i.e. **nP** = # defectives/# of observations), then calculate the average defective rate $(\mathbf{n\overline{P}})$ as

$$\mathbf{n\overline{P}} = \frac{\text{Total \# of all defectives}}{\text{Total \# of observations}}$$

- Calculate the average of defectives per data point $(\overline{\mathbf{P}})$.

$$\overline{\mathbf{P}} = \frac{\mathbf{n\overline{P}}}{\text{\# of units checked per data point}}$$

- Calculate sigma (*s*).

$$s = \sqrt{\mathbf{n\overline{P}} \times (1 - \overline{\mathbf{P}})}$$

- Calculate the upper control limit (**UCL**) for each data point. __Note__: If the calculated **UCL** value is higher than is feasible, set the **UCL** to the highest feasible value.

$$\mathbf{UCL} = \mathbf{n\overline{P}} + (3 \times s)$$

- Calculate the lower control limit (**LCL**) for each data point. __Note__: If the calculated value **LCL < 0**, then set the **LCL** to zero.

$$\mathbf{LCL} = \mathbf{n\overline{P}} - (3 \times s)$$

- Plot **nP**, $\mathbf{n\overline{P}}$, **UCL** and **LCL** on the same graph.

Figure 32. Calculations for NP chart.

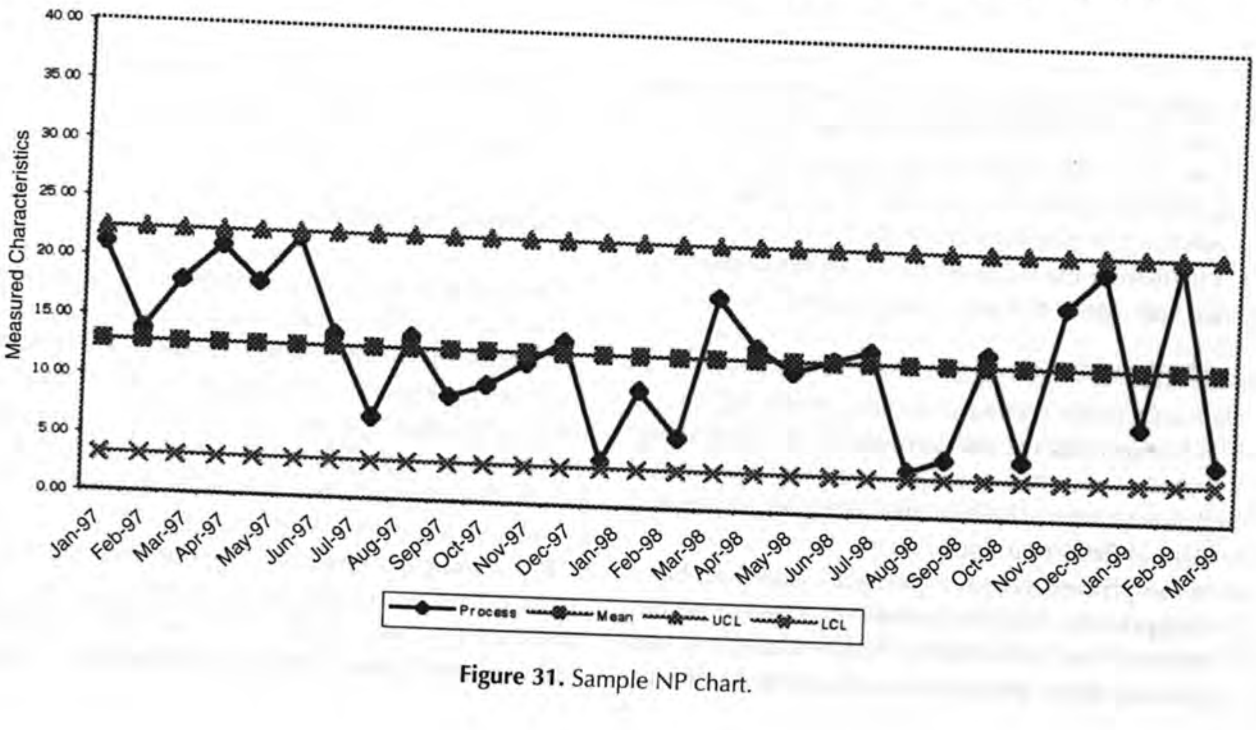


Figure 31. Sample NP chart.

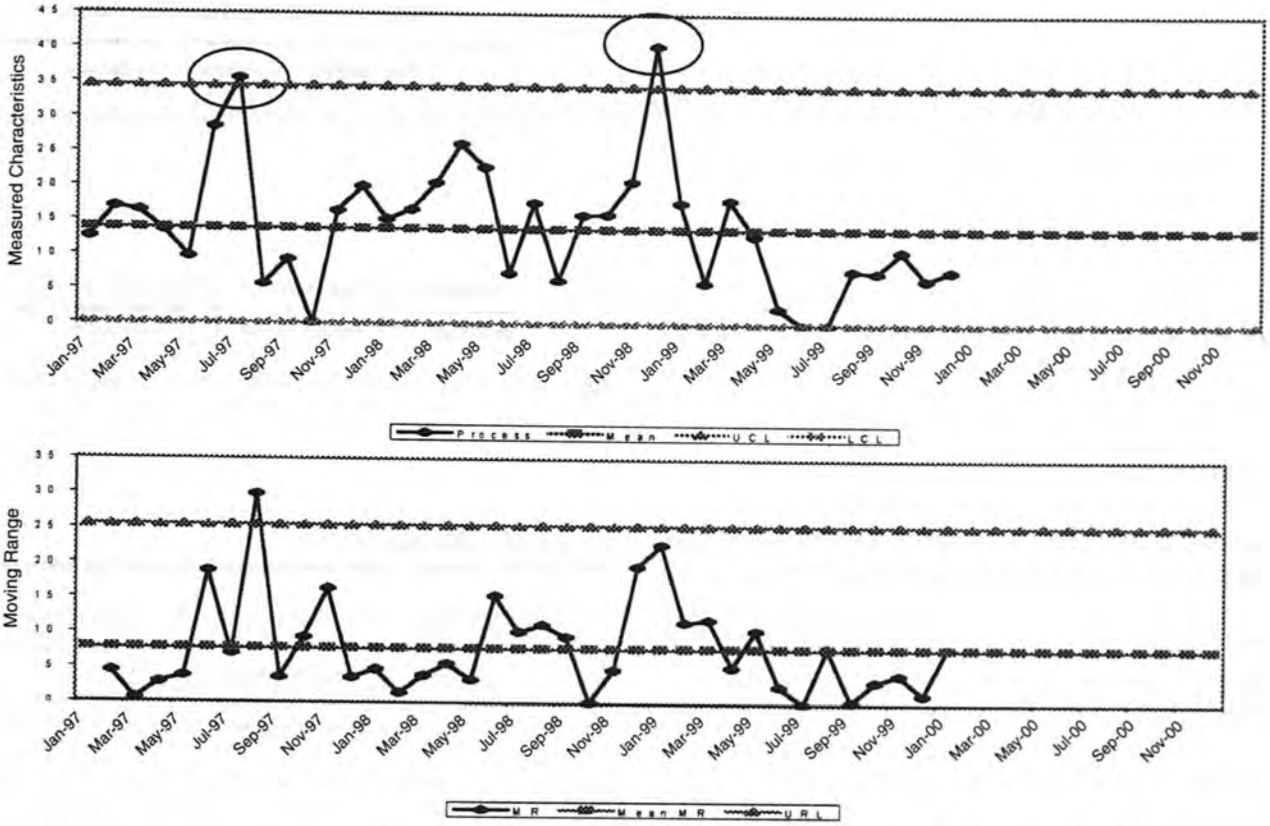


Figure 33. Data display in XmR chart.

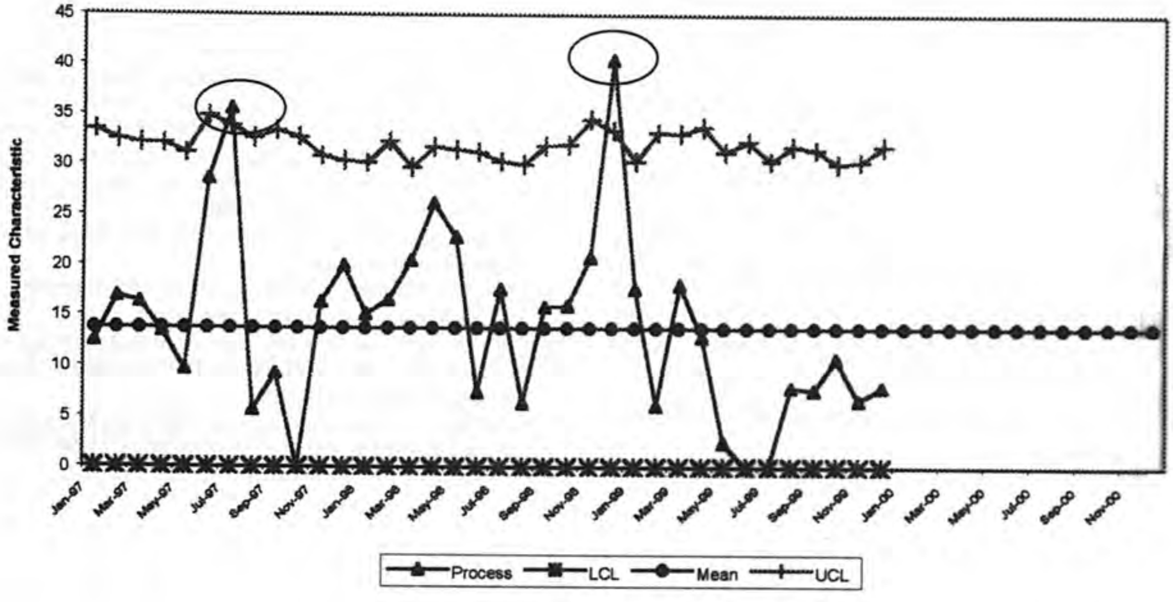


Figure 34. Data display in U chart.

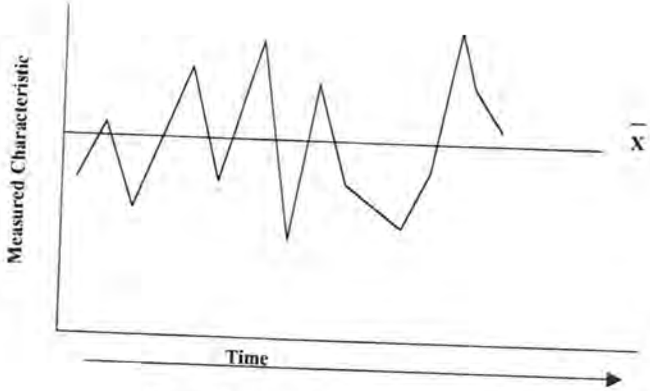


Figure 35. Run chart components.

$$\overline{X} = \frac{\text{Value}_1 + \text{Value}_2 + \text{Value}_3 + \ldots \text{Value}_n}{\text{Total number of values in the sample}}$$

Figure 36. Calculations for the center line when the value is a count.

$$\overline{X} = \frac{\text{Numerator}_1 + \text{Numerator}_2 + \ldots \text{Numerator}_n}{\text{Denominator}_1 + \text{Denominator}_2 + \ldots \text{Denominator}_n}$$

Figure 37. Calculations for the center line when the value is a rate.

files may be of significant importance when an organization first is beginning to use control charts. Typically, early on in the process, there may be only a small group of individuals who understand control charts well enough to create them. It may be advantageous for this group to set up the data and control chart graphs and to provide the files to the departments for ongoing data entry and display. One of the disadvantages of spreadsheet software is that the control chart creator needs a basic understanding of the spreadsheet software in order to set up the data, enter the formulas in the cells, and create the graphs. Additionally, understanding some of the more advanced concepts of this software can be helpful, such as the use of "if" statements to set the LCL to zero if the calculated value is less than zero.

Specially developed control chart software eliminates the need to know and understand the formulas for each type of chart because this type of software typically creates the graphs automatically when the data are entered into the program. On the downside, there is a cost associated with the purchase of these programs and electronic versions of the charts (i.e., data files) can be shared only with others who have the same software.

SUGGESTED READINGS

Carey RG and Lloyd RC. *Measuring Quality Improvement in Healthcare: A Guide to Statistical Process Control Applications*. New York, NY: Quality Resources, 1995

Hulley SB and Cummings SR. *Designing Clinical Research: An Epidemiologic Approach*. Baltimore MD: Williams & Wilkins, 1988.

Kelley DL. *How to Use Control Charts for Healthcare*. Milwaukee, WI: ASQ Quality Press, 1999.

Nelson EC, Splaine ME, Batalden PB, and Plume SK. "Building Measurement and Data Collection into Medical Practice," *Annals of Internal Medicine*, Vol. 128; No. 6, 15 March 1999, pp. 460–466.

Nolan TW and Provost LR. "Understanding Variation," *Quality Progress*, 23(5): 70–78. May 1990.

Plsek PE. "Tutorial: Introduction to Control Charts," *Quality Management in Health Care*, Vol. 1, No. 1, 1992, pp. 65–74.

Scholtes PR, Joiner BL and Streibel BJ. *The Team Handbook*. Second Edition. Madison WI: Oriel Inc., 1996.

Shewhart WJ. *Statistical Method from the Viewpoint of Quality Control*. Dover Publications, 1986.

StatSoft, Inc. "Quality Control Charts." http://www.statsoftinc.com/textbook/stquacon. html.

Sytsma S. "Common Control Chart Cookbook" http://www.sytsma.com/tqmtools/charts. html.

Wheeler DJ and Chambers DS. *Understanding Statistical Process Control*. Knoxville, TN: SPC Press, 1992. "Chapter 3: Shewhart's Control Charts" pp. 37–52.

Wheeler DJ. *Understanding Variation: The Key to Managing Chaos*, Second Edition. Knoxville, TN: SPC Press, 2000.

Appendix 1

Practice Exercise—Identifying Data Types

Instructions

For each example, identify whether it is variable data (i.e., measured) or attributes data (i.e., either a count or derived from a count) and put a ✓ in the appropriate column. Answers and explanations for the answers can be found in Appendix 4.

Data example	Variables data (measurement)	Attributes data (count or derived from a count)
1. Number of days to wait for a routine office visit appointment		
2. Medication errors resulting in death or severe injury		
3. Percent of patients in restraints		
4. Average overall satisfaction score		
5. Percent of patients in a restraint for an appropriate reason		
6. Rate of ventilator-associated pneumonia		
7. ALOS for patients undergoing a C-section		
8. Daily blood sugar results for a diabetic patient		
9. Average severity score for medication errors		
10. Percent of patients on Medicaid		

Appendix 2

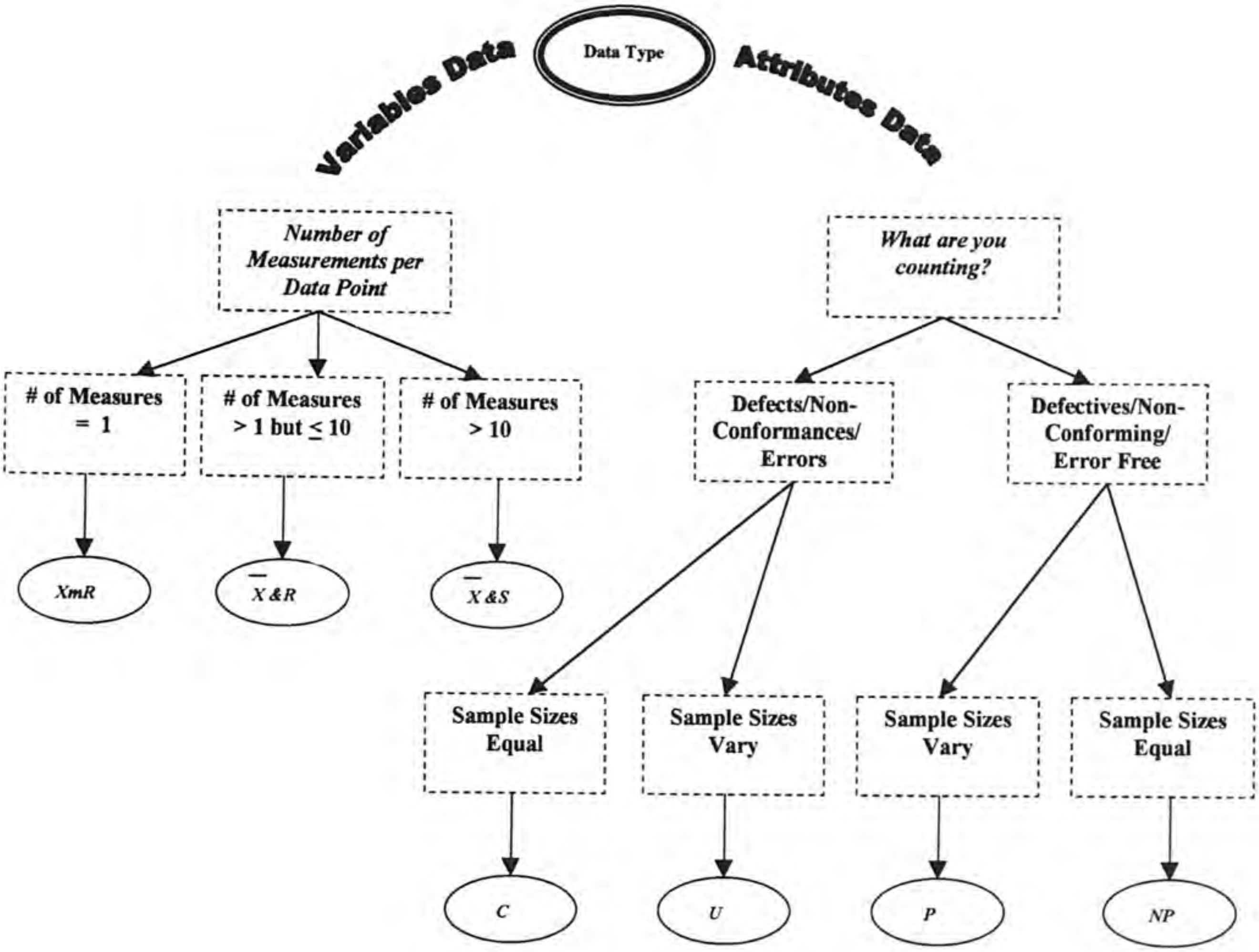
Practice Exercise—Control Chart Selection

Instructions

For each example, identify the correct control chart and put a ✓ in the appropriate column. Answers and explanations for the answers can be found in Appendix 5.

Data example	Variables charts			Attributes charts			
	XmR	$\overline{X}$ & R	$\overline{X}$ & S	C	U	P	NP
1. Rates of patient falls per month							
2. Monthly staff turnover rates							
3. Monthly budget variances							
4. Monthly primary C-section rates							
5. Number of employee needlestick injuries per month							
6. Monthly results of a satisfaction questionnaire that rates satisfaction on a scale from 1 to 5 and is given to 25 randomly selected employees each month							
7. Monthly telephone abandonment rates that are based on data for all calls to a particular number							
8. Total number of procedures performed each month by procedure type							
9. Average time to clean a patient care room after discharge based on a random sample of 5 discharges per day							
10. Monthly rates of medication errors							

Appendix 3
Control Chart Selection Process



Appendix 4
Practice Exercise Answers

Identifying Data Types (Appendix 1)

Data	Variables Data	Attributes Data	Rationale
1. Number of days to wait for a routine office visit appointment	✓		Represents a *measure* of the time it takes from when the appointment is made until when it takes place
2. Number of medication errors resulting in death or severe injury		✓	Represents a *count* of an event (i.e., medication errors resulting in death or severe injury).
3. Percent of patients in restraints		✓	Derived from a *count* of patients in restraints.
4. Average overall satisfaction score	✓		*Measure* satisfaction on a continuous scale (e.g., poor to excellent).
5. Percent of patients in a restraint for an appropriate reason		✓	Derived from a *count* of patients in restraints.
6. Rate of ventilator-associated pneumonia		✓	Derived from a *count* of patients with ventilator-associated pneumonia.
7. ALOS for patients undergoing a C-section	✓		Length of stay is a *measure* of the time from admission until discharge.
8. Daily blood sugar results for a diabetic patient	✓		Laboratory tests are *measures* of a particular value.
9. Average severity score for medication errors	✓		In this context, severity scores would be a *measure* (note difference in wording from example #2).
10. Percent of patients on Medicaid		✓	This rate is derived from a *count* of Medicaid patients.

Appendix 5
Practice Exercise Answers

Control Chart Selection (Appendix 2)

Data example	Control chart	Rationale
1. Rates of patient falls per month	**U chart** **or** **XmR**	• Fall rates are derived from a count of falls per month and thus would be considered attributes data. Because a fall could happen more than once to a patient, this is counting defects. Because the # of patients at risk varies from month to month, a **U chart** would be appropriate. • An **XmR chart** also could be used if desired (see sections on "Types of Data" and "Can More Than One Chart Be Correct" for further explanation as to rationale).
2. Monthly staff turnover rates	**P chart** **or** **NP chart** **or** **XmR**	• Turnover rates are derived from counts of staff who leave a job and thus are attributes data. Because an employee only can leave the job one time, it would be a count of defectives. If the population at risk (i.e., # of employees on role) changes from month to month (likely in a large organization), a **P chart** would be used, but if the # of employees on role is stable from month to month, an **NP chart** could be used. • An **XmR chart** also could be used if desired (see sections on "Types of Data" and "Can More Than One Chart Be Correct" for further explanation as to rationale).
3. Monthly budget variances	**XmR**	• Budget variances are derived from the measurement of dollar amount over or under budget and thus are variables data. This figure would be a single measure each month, making an **XmR chart** appropriate.
4. Monthly primary C-section rates	**P chart** **or** **XmR**	• C-section rates are derived from counts of patients who have a C-section and thus are attributes data. Because a patient either does or does not have a C-section (i.e., it could not occur

Data example	Control chart	Rationale
		more than once per pregnancy), this is counting defectives. Because the # of deliveries varies month to month, a **P chart** is appropriate. • An **XmR chart** also could be used if desired (see sections on "Types of Data" and "Can More Than One Chart Be Correct" for further explanations as to rationale).
5. Number of employee needlestick injuries per month	**C chart** **or** **U chart** **or** **XmR**	• Number of injuries represent a count and because the injury could happen to a single staff member more than once per time period, this is counting defects. If the number of staff per month that are at risk is relatively stable, then a **C chart** could be used. If the number of staff varies significantly from month to month, then a rate should be calculated and a **U chart** should be used. • An **XmR chart** also could be used if desired (see sections on "Types of Data" and "Can More Than One Chart Be Correct" for further explanation as to rationale).
6. Monthly results of a satisfaction questionnaire that rates satisfaction on a scale from 1 to 5 and is given to 25 randomly selected employees each month	**$\overline{X}$ & S**	Satisfaction scores represent measurements on a continuous scale from 1 to 5 and thus are variables data. Because each data point is made up of 25 subvalues, an **$\overline{X}$ & S chart** is appropriate.
7. Monthly telephone abandonment rates that are based on data for all calls to a particular number	**P chart** **or** **XmR**	• Abandonment rates are derived from counts of phone calls that are abandoned and thus are attributes data. An individual call can be abandoned only once and thus would be a defective and not a defect. Because the volume of calls varies from month to month, a **P chart** would be appropriate. • An **XmR chart** also could be used if desired (see sections on "Types of Data" and "Can More Than One Chart Be Correct" for further explanation as to rationale).

Data example	Control chart	Rationale
8. Total number of procedures performed each month by procedure type	XmR	• Even though total number of procedures sounds like a count, it is not a count of defects or defectives. In reality, it is a measurement on a continuous scale (i.e., from 0 to some maximum number of procedures). Consequently, it would be considered to be variables data. Each data point would represent a single measurement and thus an **XmR chart** would be the appropriate chart.
9. Average time to clean a patient care room after discharge based on a random sample of 5 discharges per day	$\bar{X}$ & R chart	Time to clean a room would be a measurement from the time it was started until the time that it was completed. Because each data point represents a sample of 5 measurements, an **$\bar{X}$ & R chart** would be appropriate.
10. Monthly rates of medication errors	U chart or XmR	• Medication error rates are derived from counts of medication errors and thus are attributes data. Because medication errors can occur more than once to an individual during the time period or multiple errors could be made with a single dose of medication, medication errors would be considered defects. The population at risk (e.g., patients at risk per month or medication doses dispensed per month) varies each time period and thus a **U chart** should be used. • An **XmR chart** also could be used if desired (see sections on "Types of Data" and "Can More Than One Chart Be Correct" for further explanation as to rationale).