Chapter 5

Choosing Our Own Rules

To ensure that our methods meet defined quality specifications, we must determine the best QC rules or limits to use. Appropriate limits for accepting or rejecting a run should be based on current method accuracy and precision as reflected by the calculated mean and SD, and chosen to alert us to changes that exceed quality specifications defined by the target value and the total error allowable (TEa) for each control.

BACKGROUND

Medical laboratories have been doing QC for more than five decades. In many instances the process has changed little since Levey and Jennings' publication in 1950. The QC rules we use, the action limits on our QC charts, the number of controls, and the frequency with which we run controls are often based on historical practice within our institution or on outdated recommendations. Yet selecting a QC process without regard for quality goals or requirements is, at best, "arbitrary."

We are all trying to achieve the best possible accuracy and precision. We often want our control results to stay within our mean ± 2 SD. The problem is that the mean and SD on the same control material for the same test frequently differ significantly between laboratories.

This chapter examines the concept of QC design, its variables and its applications, and examples of QC systems.

WHAT IS QC DESIGN?

Do traditional QC systems really "control" quality? The idea of "control" implies that something or someone can impact results. The speedometer in your car monitors speed. It can tell you if you are going too fast (out of control), but it does not change the speed of the car. As the driver (controller) of the car, you have an impact on how fast you go. You control the speed of the car. If you are going too fast (out of control) you can ease up on the gas pedal and slow the car down. You are now in control. You achieved your goal by exercising your control over the system.

In this sense, a traditional monitoring system that does little more than put dots on QC charts and respond to a value outside some arbitrary limit (e.g., ±2 SD) is like having a speedometer with numbers from −3 through 0 to +3. When the needle hits plus +2 or −2, the driver would slow down or speed up. But the "2 SD" limit on the speedometer could be 150 in one lab and 50 in another. They are using different speedometers, yet, in real terms, both drivers are required to stay

within the same "speed limits." Similarly, two or more laboratories must meet the same clinical and PT requirements. A QC chart that merely warns us when our values are outside ± 2 SD does not ensure that our laboratory will meet the required quality specifications.

How does Performance-Driven Quality Control (PDQC) differ from arbitrary quality control? A properly labeled QC chart using the actual mean and SD and appropriate control limits alerts us to changes in method performance. We usually consider results acceptable if they are within the control limits, and unacceptable if they are outside the specified limits or break a QC rule. As we have mentioned, these SD limits vary significantly from lab to lab—as do the control limits or QC rules in use.

When method performance changes, QC flags on the QC chart usually increase. For example, this can occur when a new population of data changes the mean and/or SD. Although QC flags alert us to changes in the analytical system, they do not tell us whether or not we are still meeting performance specifications, or why a specific control result is outside QC limits.

As noted in Chapter 4, the probability of detecting true errors or significant change is affected by the assigned mean and SD used on the QC chart. "Arbitrary" QC often sets the mean or SD at values that do not represent current actual method performance. In an arbitrary QC system, the same QC rules are usually applied to all controls for all tests.

In contrast, Performance-Driven Quality Control (see Figure 5-1) lets us assess the summary statistics from daily QC results (mean and SD) to evaluate method performance against defined quality specifications (target and TEa). We then use the actual current mean and SD on the QC chart and select appropriate QC rules to maximize error detection and minimize false rejection while maintaining each control within its defined quality specification.

Do arbitrary QC charts reflect method performance? We use the term "arbitrary quality control" to refer to the process where labs assign the mean and SD on their QC charts and select QC rules, usually including the 1-2s rule, without considering quality specifications. The mean and SD might be based on cumulative performance for each control. These numbers might arbitrarily be assigned based on recommendations from instrument manufacturers, control manufacturers, or even performance goals provided in the literature, or from proficiency testing programs. Arbitrarily selected limits and rules often bear no relationship to quality specifications. QC results that are acceptable based on arbitrarily designed QC charts will not necessarily meet performance requirements.

Figures 5-2 and 5-3 illustrate two QC charts for different sodium methods. The QC chart in Figure 5-2 has a mean of 142 and SD of 1.0. The QC chart in Figure 5-3 illustrates a laboratory with a mean of 142 and SD of 2.0. Notice that the QC charts appear identical, except for the assigned SD values.

The Z-Bars© in Figure 5-4 show the relative performance of these methods when compared to a target value of 142.0 mmol/L and TEa limits of ± 4.0

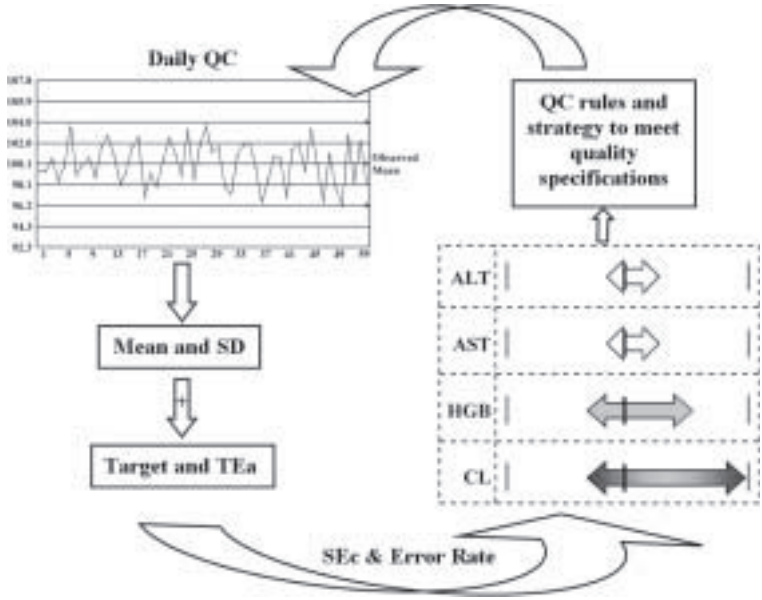


Figure 5-1. Summary statistics from control results are compared to quality specifications to determine requirements for QC rules and strategy.

mmol/L as recommended by CLIA. Although the QC charts look similar, the level of performance in these two methods differs. By comparing the mean and SD to the quality specifications for this control, we easily see that method A, with an SD of 1.0, has a far better probability of passing proficiency surveys and of meeting clinical requirements than does method B.

Why should we design a QC system? Most laboratories have embraced technical advances in instruments and methods, but few have modified their QC pro-

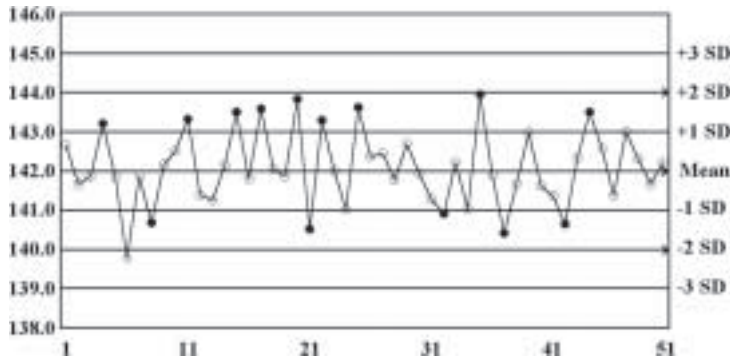


Figure 5-2. Sodium QC chart: mean, 142, and SD, 1.0.

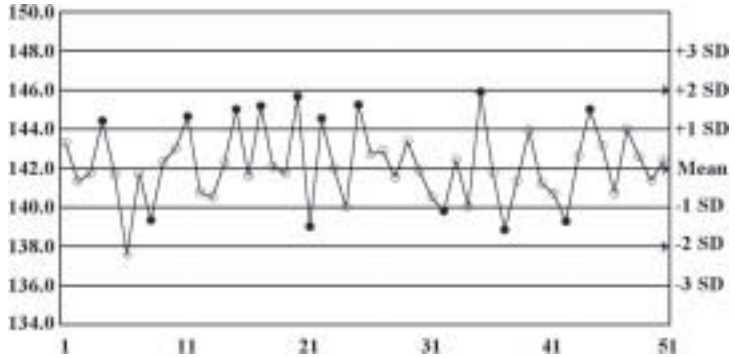


Figure 5-3. Sodium QC chart: mean, 142, and SD, 2.0.

cesses to match: many laboratories are still using a 1-2s rule as recommended in 1950. This approach worked well for manual methodologies, but think of the changes in the technology since then!

With the dramatically improved precision of today's methods, it is not uncommon to see shifts of several SD for the same control on the same test from time to time within a single laboratory, particularly when changing reagent lot numbers. For the past 10 years leading clinical biochemists such as James Westgard, Per Hyltoft Petersen, and others have advised using a variety of QC rules to match the analytical capability and stability of each method, relative to a defined quality specification.

Performance-Driven Quality Control helps us design QC systems that usually alert us quickly to significant changes, and that generate few QC flags when the system is operating safely within quality specifications.

How do we assess the maximum acceptable shift in the mean? In a poster presented at the 2000 AACC annual meeting, my co-workers and I studied critical

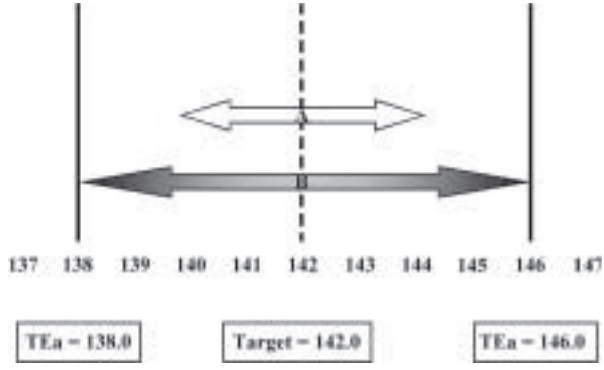


Figure 5-4. Z-Bars of sodium controls with SD of 1.0 and 2.0 compared to TEa of 4.0.

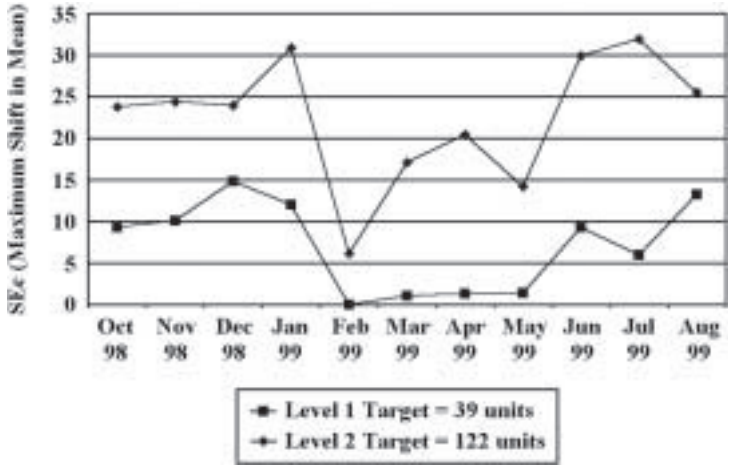


Figure 5-5. Variation in ΔSEc for AST on a single analyzer over 11 months.

systematic error (ΔSEc) on monthly QC summary data to evaluate the range of observed ΔSEc values and the efficacy of modifying QC rules based on ΔSEc and method stability. We analyzed output from Quality Advisor© software[1], from six Vitros 250 analyzers[2] in six hospital laboratories from October 1998 to August 1999. Target values (based on peer comparison data) and TEa limits (specified empirically by the laboratory director to meet clinical requirements) were identified for each level of each control for each test on each instrument. Mean and SD values, generated each month from routine daily QC data, were entered in real time into Quality Advisor©. The software calculated ΔSEc as

$$[(TEa - |Bias|)/SD] - 1.65$$

and recommended QC rules based on ΔSEc and method stability.

This study examined ΔSEc values and QC rule recommendations on 2712 QC summary results from 23 routine chemistry tests in six hospital laboratories over 11 months. We exported data from Quality Advisor to Microsoft Excel and SPSS to create summary graphs. Figure 5-5 illustrates changes in ΔSEc for two levels of control for AST on the same analyzer over an 11-month period. Each point shows the critical systematic error (ΔSEc): the number of SD the mean could shift before the probability that this method would produce results that could be clinically misleading or fail a proficiency survey was greater that 5%.

In the first 4 months, the mean could shift more than 10 SD. In February 1999, perhaps with a new reagent lot, the lower level control was at (or exceeding!) the error limit. It stayed close to the TEa limit for 4 months, then returned to perfor-

[1] © Q.I.K. Quality Is Key Ltd., Worthington, Ontario, Canada
[2] Ortho Clinical Diagnostics, Raritan, New Jersey, USA

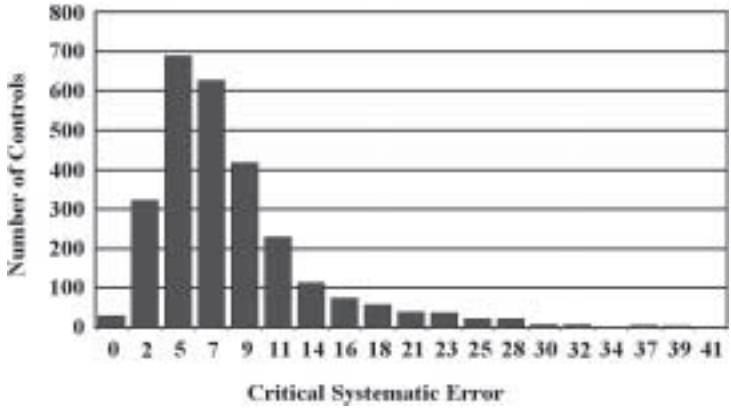


Figure 5-6. ΔSEc values calculated on 2,712 QC summary results from 23 routine chemistry tests in six hospital laboratories over 11 months.

mance levels that were well within the quality specifications. (The lower-level control had a target value of 39 units, and TEa limit of 6 14.5 units. The higher-level control had a target value of 122 units and TEa of ± 27 units.)

When we compare method performance to defined quality specifications (rather than to last month's data), we can design strategies that will warn us when QC data points exceed acceptable performance—with a minimal number of false QC flags. Figure 5-6 shows the ΔSEc values calculated on 2,712 QC summary results from 23 routine chemistry tests in six hospital laboratories over 11 months. Notice that most of the controls could shift well over three SD before exceeding quality specifications.

When methods are close to the TEa limit (have a low ΔSEc), or when methods have a high expected error rate, we want to design our QC system to improve error detection. In these cases, we will select a higher number of controls or run our existing controls more frequently, and we will select QC rules that are more powerful when assessing small changes in method performance (maximizing true rejects). When the ΔSEc is high (method performance is well within quality specifications) and the analytical system has a low expected error rate, we will run fewer controls, or run controls less frequently, and we will select QC rules to minimize false rejects of false-positive QC flags.

VARIABLES IN QC DESIGN

How does the QC process interact with the analytical process? Two processes must be in place before reporting a patient result to the clinician or a PT result to the organizers of the proficiency-testing program. As illustrated in Figure 5-7, when we design our QC system, we "balance" the QC system to meet the changing performance and stability of the analytical system.

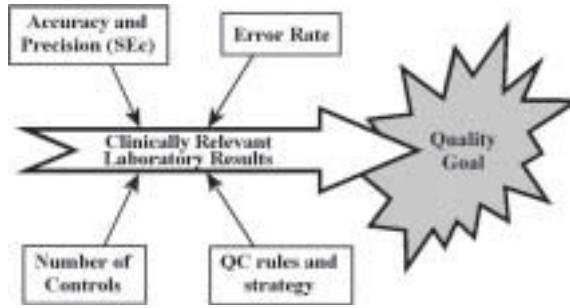


Figure 5-7. QC rules, number of controls, and other strategies vary to match changes in analytical performance.

Our analytical processes vary in their susceptibility to the occurrence of significant errors. Error rates vary from low to moderate to high, depending on the frequency of significant errors that occur in a specific analytical system. Some methods seldom encounter problems, while others are susceptible to relatively frequent sources of significant error. Methods with high error rates frequently demonstrate significant shifts in the mean or ongoing precision problems; these methods may also be susceptible to frequent instrument breakdowns or other performance problems.

We monitor method performance (accuracy and precision) relative to a quality specification by calculating critical systematic error (ΔSEc). ΔSEc is an extremely powerful and useful statistic. ΔSEc indicates in one number how method accuracy and precision compare to the target and TEa limit set for each control. ΔSEc indicates the number of SD the mean can shift before more than 5% of results exceed TEa limits. Changes in either the mean or SD will change critical systematic error.

How QC rules affect QC design. A QC rule combines a limit, and the number of control values that need to exceed that limit for a run to be judged out of control. For example, the 1-3s rule requires one point to exceed 3 SD for the run to be labeled "out." The 4-1s requires four points to exceed 1 SD to generate a QC flag. We see a 1-2s QC flag when a single result is more than 2 SD from the mean, and we see a 1-3.5s QC flag if a single result is more than 3.5 SD from the mean. A 1-3s or a 1-3.5s rule will only generate a QC flag when method performance significantly changes. Remember that virtually 100% of our data will fall within ± 3 SD of the mean, as long as no change occurs in our system. The 10-x rule (which detects a very small shift in the mean) is violated when 10 results fall on the same side of the mean. Other rules detect different sizes and types of change, for example, systematic change, increased random error, etc.

James Westgard and others have performed computer simulations to predict the probability of error detection and false rejection for a variety of QC rules and multi-rule combinations. Figure 5-8 illustrates how shifts of 1, 2, 3, or 4 SD in the

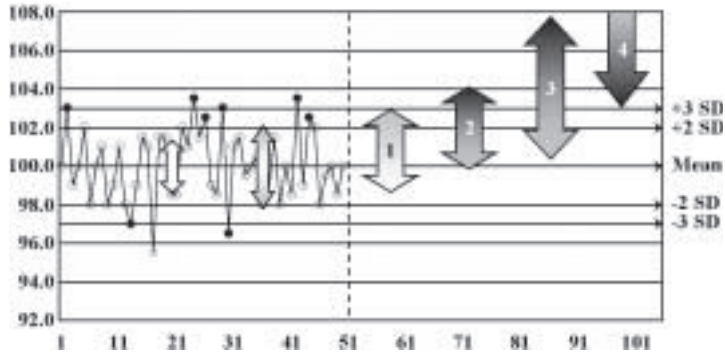


Figure 5-8. The probability of detecting a shift in the mean is proportional to the size of the shift that occurs.

mean will cause varying proportions of the new data population to exceed the ± 2 SD or ± 3 SD limits. Notice that after a shift in the mean of +1 SD, most of the new data population will remain within the previously defined ± 2 SD limits. If the mean shifts 2 SD, half of the new data will exceed ± 2 SD. As the size of the shift in the mean increases, the portion of data outside the assigned ± 2 SD or ± 3 SD limits increases. The probability of detecting a shift in the mean is therefore proportional to the size of the shift that occurs.

Table 5-1 shows the probability of error detection (Ped) and probability of false rejection (Pfr) with common QC rules, with N (the number of control samples in each run) of 2 and 3. When there is no change, the shift in the mean is 0. Notice how the various rules differ in Ped and Pfr as the mean shifts. In this set of common QC rules, the 1-2s rule has the highest probability of detecting a shift in the mean of 1 SD. When the ΔSEc for a method is very low, it may be appropriate

Table 5-1. Probability of error detection and false rejection with common QC rules.

			Probability of QC flag with mean shift				
QC Rule	**N**	**# of Runs**	**0 SD**	**1 SD**	**2 SD**	**3 SD**	**4 SD**
1-2s	2	1	10%	30%	76%	98%	100%
1-2.5s	2	1	3%	14%	55%	90%	100%
1-3s	2	1	0%	5%	32%	75%	97%
1-3.5s	2	1	0%	2%	15%	49%	88%
1-3s/2-2s/R-4s/4-1s/10x	2	5	0%	25%	90%	100%	100%
1-2s	3	1	15%	34%	84%	98%	100%
1-2.5s	3	1	4%	17%	62%	95%	100%
1-3s	3	1	1%	5%	38%	83%	99%
1-3.5s	3	1	0%	2%	15%	58%	95%
1-3s/2-2s/R-4s/3-1s/12x	3	4	0%	38%	94%	100%	100%

to use the 1-2s rule as a reject signal. Note, however, the high price we pay in false rejects.

The multi-rule combinations appear to provide the lowest false rejection (Pfr) and highest error detection (Ped); however these values are based on 4 or 5 QC runs. The 4-1s and 10x or 12x rules, often used as warnings rather than reject signals, require multiple runs before they can generate a QC flag. How many patient samples might we report before we see a QC flag?

When the mean can shift 4 SD, most of the QC rules or combinations have a high probability of detecting this change in the first QC run. By calculating ΔSEc, we know the size of shift we must detect to maintain acceptable method performance. We can then choose appropriate QC rules to maximize Ped and minimize Pfr.

How critical systematic error (ΔSEc) affects QC design. When we know the size of the shift that we need to detect and the relative probability of errors (ΔSEc and error rate), we can choose appropriate QC rules to maximize the chances that we will detect errors before our methods exceed quality specifications. Figure 5-9 illustrates four examples with varying acceptable shifts in the mean before exceeding the specified TEa limit for each control. We can easily visualize how it would be appropriate to use "tighter" QC rules when the method is closer to the error limit, and to use "looser" QC rules when a method can shift many standard deviations before exceeding the quality specification.

How "N" affects QC design. "N" describes the number of controls in a run of samples. When N = 2 there are two levels of control or two replicates of a single control in each analytical run; when N=3 there are 3 controls in each run. To design a QC process that can detect a very small change in the analytical system, increase the number or frequency of the controls.

When we run more levels of controls, or run the same controls more frequently, we increase our probability of detecting errors. When we increase the

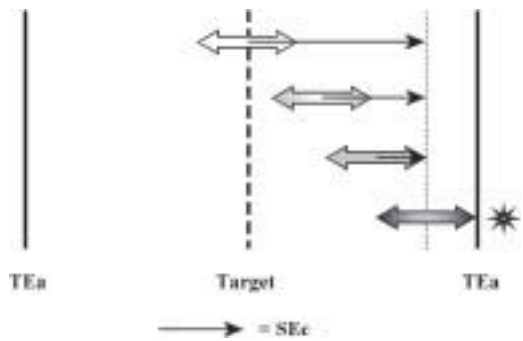


Figure 5-9. ΔSEc measures the maximum acceptable shift in the mean before results exceed quality specifications.

number of controls, we increase the probability of error detection. Unfortunately, we also increase the probability of false rejection. Remember, if we are using a 1-2S rule, we will see 5% of our "good" data from each control falling between 2 and 3 SD. Therefore, the more controls we run, the higher the probability in any given run that one of them will fall outside 2 SD.

APPLYING QC DESIGN

PDQC enables the selection of appropriate QC rules and strategies to maintain performance of each analytical process within defined quality specifications (see Table 5-2).

QIK-QC Strategy Tables© help us meet quality goals: our objective is to design a QC system that will maximize true rejects and minimize false rejects. By customizing the QC process to meet defined quality requirements, virtually every QC flag will be a significant event.

These tables help us select appropriate QC rules and determine any additional steps we must take to improve error detection or initiate corrective action. We select QIK QC Strategy Tables© based on the error rate of each method and the number of controls we routinely use for each test (2 or 3 levels of controls). Table columns indicate the critical systematic error calculated for a control, the appropriate selection for a single QC rule, and the appropriate selection for a multi-rule combination. Symbols on the strategy table indicate additional QC strategies that we may apply to increase error detection, such as

- examining QC charts more frequently, as we will often notice small shifts in the mean on the QC chart before we will react to statistical QC flags;
- increasing control frequency or N to improve the probability of error detection; or

Table 5-2. Steps for Performance-Driven Quality Control

1. Specify the target value for each control for each test.
2. Specify the total error allowable about the target value (TEa).
3. Categorize the method error rate as low, moderate or high.
4. Calculate the current method mean.
5. Calculate the current method SD, or specify the "usual" SD based on historical performance.
6. Calculate critical systematic error (ΔSEc).
7. Choose the appropriate QIK–QC Strategy Table© based on the number of controls and error rate.
8. Select the appropriate QC strategy based on ΔSEc, including:
   - 8.1. Single QC rule or multi-rule combination
   - 8.2. Control frequency
   - 8.3. Frequency of visual examination of QC charts
   - 8.4. Necessity for corrective action or method improvement

- if the error rate is high or ΔSEc is low, initiating corrective action to improve the accuracy, precision, or stability of this method.

Six QIK-QC Strategy Tables© are provided for N of 2 or 3, and error rates of low, moderate, and high. The recommended QC rules and strategy vary on each table with the ΔSEc of each control:

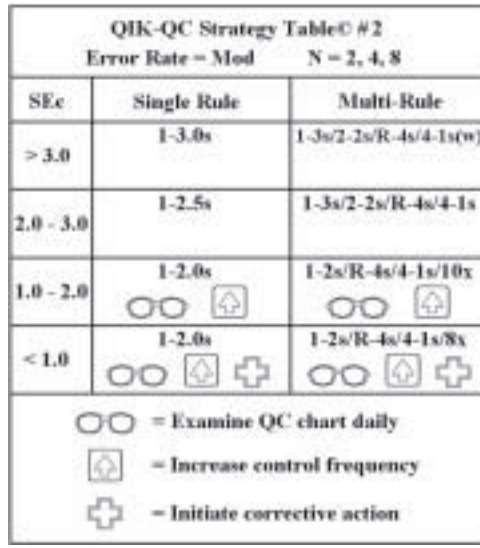
- If ΔSEc is high (>3), choose QC rules to minimize Pfr.
- If ΔSEc is between 2 and 3, choose QC rules to increase Ped.
- If ΔSEc is < 2, choose QC rules to increase Ped, increase N, and increase visual examination of QC charts.
- If ΔSEc is < 1, choose QC rules to maximize Ped, increase N, increase visual examination of QC charts, and initiate corrective action.

The strategy tables provided in Figures 5-10 to 5-15 illustrate one approach to Performance-Driven Quality Control. This template may be modified to vary ΔSEc breakpoints, error rate classifications, and recommended QC rules and strategy to match changing needs and preferences of individual laboratory professionals. Computer software programs such as Quality Advisor© simplify the customization of QC design.

How ΔSEc and error rate affect QC design. The current performance of a control sample for any analytical method is defined by two terms: critical systematic error and error rate. ΔSEc varies from 0 to >3 with differing quality specifications, method accuracy, and precision. A high ΔSEc indicates that a control can

QIK-QC Strategy Table© #1 Error Rate = Low N = 2, 4, 8		
SEc	Single Rule	Multi-Rule
> 3.0	1-3.5s	1-3s/4-1s(w)
2.0 - 3.0	1-3.0s	1-3s/2-2s/R-4s/4-1s(w)
1.0 - 2.0	1-2.5s OO	1-3s/2-2s/R-4s/4-1s OO
< 1.0	1-2.0s OO ✚	1-2s/R-4s/4-1s/10x OO ✚
OO = Examine QC chart daily ⌂ = Increase control frequency ✚ = Initiate corrective action		

Figure 5-10. QIK-QC Strategy Tables© #1 N = 2, 4, 8. Error Rate = Low.

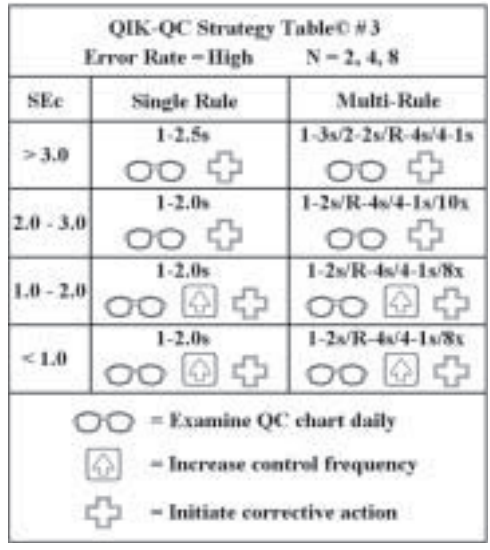


QIK-QC Strategy Table© # 2
Error Rate = Mod N = 2, 4, 8

SEc	Single Rule	Multi-Rule
> 3.0	1-3.0s	1-3s/2-2s/R-4s/4-1s(w)
2.0 - 3.0	1-2.5s	1-3s/2-2s/R-4s/4-1s
1.0 - 2.0	1-2.0s	1-2s/R-4s/4-1s/10x
< 1.0	1-2.0s	1-2s/R-4s/4-1s/8x

= Examine QC chart daily
= Increase control frequency
= Initiate corrective action

Figure 5-11. QIK-QC Strategy Tables© #2 N = 2, 4, 8. Error Rate = Moderate.

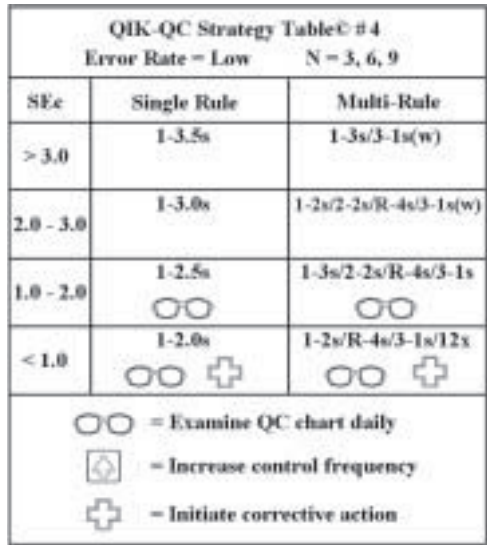


QIK-QC Strategy Table© # 3
Error Rate = High N = 2, 4, 8

SEc	Single Rule	Multi-Rule
> 3.0	1-2.5s	1-3s/2-2s/R-4s/4-1s
2.0 - 3.0	1-2.0s	1-2s/R-4s/4-1s/10x
1.0 - 2.0	1-2.0s	1-2s/R-4s/4-1s/8x
< 1.0	1-2.0s	1-2s/R-4s/4-1s/8x

= Examine QC chart daily
= Increase control frequency
= Initiate corrective action

Figure 5-12. QIK-QC Strategy Tables© #3 N = 2, 4, 8. Error Rate = High.

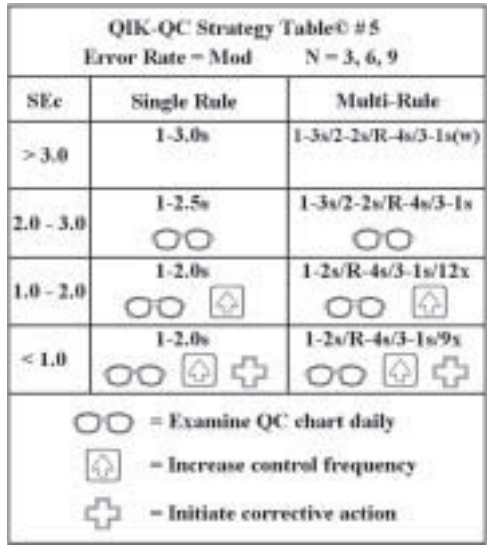


QIK-QC Strategy Table© #4
Error Rate = Low N = 3, 6, 9

SEc	Single Rule	Multi-Rule
> 3.0	1-3.5s	1-3s/3-1s(w)
2.0 - 3.0	1-3.0s	1-2s/2-2s/R-4s/3-1s(w)
1.0 - 2.0	1-2.5s	1-3s/2-2s/R-4s/3-1s
< 1.0	1-2.0s	1-2s/R-4s/3-1s/12x

= Examine QC chart daily
= Increase control frequency
= Initiate corrective action

Figure 5-13. QIK-QC Strategy Table© #4 N = 3, 6, 9 Error Rate = Low



QIK-QC Strategy Table© #5
Error Rate = Mod N = 3, 6, 9

SEc	Single Rule	Multi-Rule
> 3.0	1-3.0s	1-3s/2-2s/R-4s/3-1s(w)
2.0 - 3.0	1-2.5s	1-3s/2-2s/R-4s/3-1s
1.0 - 2.0	1-2.0s	1-2s/R-4s/3-1s/12x
< 1.0	1-2.0s	1-2s/R-4s/3-1s/9x

= Examine QC chart daily
= Increase control frequency
= Initiate corrective action

Figure 5-14. QIK-QC Strategy Table© #5 N = 3, 6, 9 Error Rate = Moderate

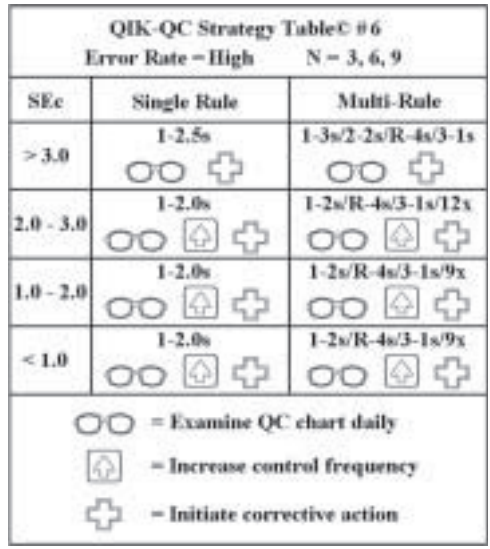


Figure 5-15. QIK-QC Strategy Table© #6 N = 3, 6, 9 Error Rate = High

tolerate a relatively large change in accuracy or precision before exceeding quality specifications. The error rate will vary between methods, and may vary within a single method from low to high with aging instrumentation or with seasonal environmental changes within a laboratory. Error rate is set at "low," "moderate," or "high" based on the experience of laboratory staff with each analytical process, as may be recorded in a log book or software program. One approach to error rate may be to assume that

- a method that experiences significant problems twice a year would be categorized as having a low error rate;
- a method that experiences significant problems every two months would be categorized as having a moderate error rate; and
- a method that experiences significant problems every month would be categorized as having a high error rate.

The examples of error rates above are one possible approach. The classification of error rate as low, moderate, or high may vary between laboratories. Methods that are tested infrequently require different criteria. More definitive definitions of error rate will likely evolve as experience with this process grows.

Note that a shift in the mean or increase in SD is not necessarily detected as soon as it occurs: it may take several QC runs to detect a 1, 2, or even 3 SD shift in the mean or a significant increase in SD. QC rules and multi-rule combinations inherently differ. The probability of detecting a significant change in method performance increases with the magnitude of the observed change in method accuracy or precision.

Knowing the capabilities of QC rules allows us to select appropriate rules and strategy for each analytical process. We modify our QC strategy to reduce false rejects or false-positive QC flags when the error rate is low and ΔSEc of a control is high, indicating that a large shift in the mean can be tolerated. If method performance changes, we change the QC strategy accordingly.

QIK-QC Strategy Tables© N = 2, 4, 8. It is common practice in many departments to use two levels of control in a run. To increase error detection, run controls more frequently. This equates to having 4 or 8 controls in each QC run.

QIK-QC Strategy Tables© for N of 2, 4, or 8 are shown in Figures 5-10, 5-11, and 5-12.

To use these tables:

1. Select the appropriate strategy table based on method error rate: low, moderate, or high.
2. Use the row of the table that corresponds to the calculated ΔSEc.
3. Select the corresponding single rule or multi-rule combination for that row.
4. When indicated, apply additional QC strategies to increase error detection:
   a. Examine the QC chart daily.
   b. Increase control frequency.
   c. Initiate corrective action.

QIK-QC Strategy Tables© N = 3, 6, 9. For many tests, we routinely use three levels of control in a run. To increase error detection, run these levels of controls more frequently. This equates to having 6 or 9 controls in each QC run.

QC DESIGN: CHEMISTRY

This example shows a sodium test with 2 levels of control. The calculations of ΔSEc, as shown in Table 5-3, indicate that each control can shift only 0.9 SD before reaching the TEa limit of 4 mmol/L as specified by CLIA. Z-Bars© in Figure 5-16 illustrate method performance and ΔSEc for each of the controls.

QIK-QC Strategy Table© #1 applies to a method with a low error rate and N = 2, 4, 8 (see Figure 5-17). With a critical systematic error < 1, we select a 1-2s rule or a multi-rule combination of 1-2s, R-4s, 4-1s, and 10x. This single rule, or

Table 5-3. Sodium Control Results

Sodium	Control 1	Control 2
Mean	143.5	121.5
SD	1.0	1.0
Target	142.0	120.0
TEa	4.0	4.0
ΔSEc	0.9	0.9

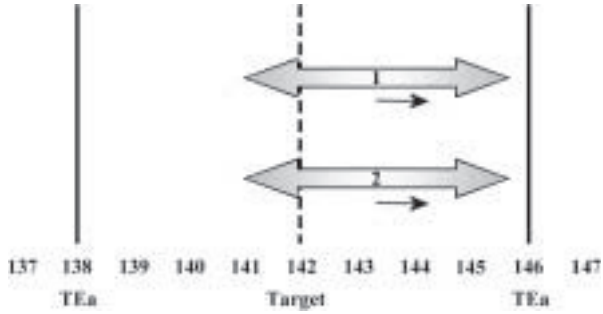


Figure 5-16. Z-Bars© illustrate method performance and ΔSEc for the sodium control results in Table 5-3.

combination of rules, will maximize our ability to detect this small shift in the mean. Note that the 1-2s rule is used in this situation as a reject signal, and not as a warning rule. A shift in the mean of < 1 SD is difficult to detect with traditional QC rules, so we would also examine the QC chart daily and initiate corrective action to attempt to improve the accuracy and precision of this method. We would also check the validity of the target value and TEa limits.

QC DESIGN: IMMUNOASSAY

This example illustrates a digoxin method with 3 levels of control. Target values are set based on peer comparison data. TEa limits are based on clinical criteria set

QIK-QC Strategy Table© #1 Error Rate = Low N = 2, 4, 8		
SEc	Single Rule	Multi-Rule
> 3.0	1-3.5s	1-3s/4-1s(w)
2.0 - 3.0	1-3.0s	1-3s/2-2s/R-4s/4-1s(w)
1.0 - 2.0	1-2.5s	1-3s/2-2s/R-4s/4-1s
< 1.0	1-2.0s	1-2s/R-4s/4-1s/10x
= Examine QC chart daily = Increase control frequency = Initiate corrective action		

Figure 5-17. QIK-QC Strategy Table© #1 recommends appropriate QC rules and strategy for a method with low error rate and controls of 2, 4, or 8.

Table 5-4. Digixin Control Results

Digoxin	Control 1	Control 2	Control 3
Mean	1.0	2.0	4.0
SD	0.1	0.2	0.2
Target	1.0	2.0	4.0
TEa	0.3	0.4	1.0
ΔSEc	2.1	0.9	2.5

by the laboratory director. As shown in Table 5-4, the ΔSEc calculated for level 1 is 2.1, level 2 is 0.9, and level 3 is 2.5.

The Z-Bars© in Figure 5-18 illustrate method performance and critical systematic error. Notice that ΔSEc varies between control levels.

Figure 5-19 illustrates the selection of QIK-QC Strategy Table© #5 for a method with a moderate error rate and N of 3. For control levels 1 and 3, we can use a 1-2.5s rule, or a combination of multi-rules of 1-3s, 2-2s, R-4s, and 3-1s. Because this method exhibits a moderate error rate, we would also want to examine our QC charts frequently (visual examination of QC charts increases the likelihood of detecting changes quickly).

For level 2, we use a single rule of 1-2s, or a multi-rule combination of 1-2s, R-4s, 3-1s, and 9x. We will examine the QC charts frequently, increase the number of controls we are using, and initiate quality improvement measures. If our daily QC system will not allow us to select individual rules for different levels of control, then we use the most demanding set of QC rules (as required for level 2) for all levels.

QC DESIGN: HEMATOLOGY

This white blood cell count example with 3 levels of control illustrates a method with a low error rate and a high ΔSEc (see Table 5-5). The TEa limits in this example are based on CLIA criteria of 15%.

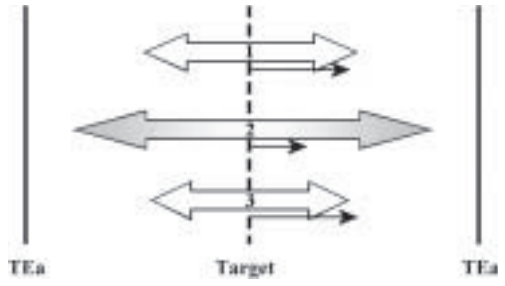


Figure 5-18. Z-Bars© illustrate method performance and ΔSEc for the digoxin control results in Table 5-4.

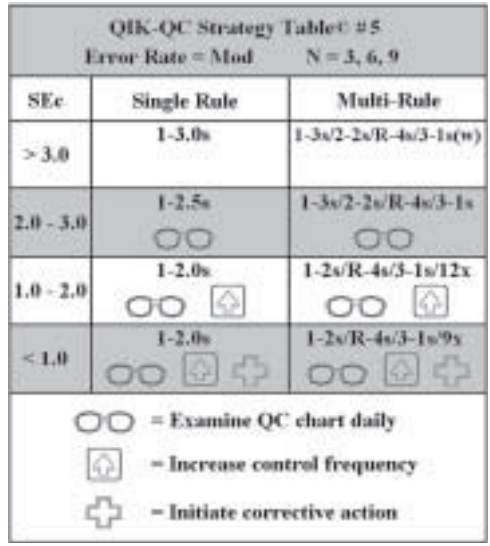


Figure 5-19. QIK-QC Strategy Table© #5 recommends appropriate QC rules and strategy for a method with a moderate error rate and controls of 3, 6, or 9.

The Z-Bars© in Figure 5-20 illustrate that these methods are performing with excellent accuracy and precision and can shift many SD before exceeding the total error limit.

We have selected the QIK-QC Strategy Table© #4, as shown in Figure 5-21, for a method with a low error rate and N of 3. As all the control levels have critical systematic errors > 3, we will select a 1-3.5s rule, or a multi-rule combination of a 1-3s with a 3-1s as a warning. Methods with a high ΔSEc and low error rate require less QC effort than those with a lower ΔSEc or higher error rate. Applying strategy tables to design QC processes allows us to "right-size" our QC process to be sure that we meet quality specifications with minimal false rejection.

QC DESIGN: COAGULATION

In this example for prothrombin time, with 2 levels of control, ΔSEc is calculated to be 1.3 for level 1 and 0.9 for level 2 (see Table 5-6). The Z-Bars© in Figure 5-22

Table 5-5. WBC Control Results

WBC	Control 1	Control 2	Control 3
Mean	2.5	10.0	20.0
SD	0.05	0.25	0.50
Target	2.5	10.0	20.0
TEa	0.38	1.5	3.0
ΔSEc	5.9	4.4	4.4

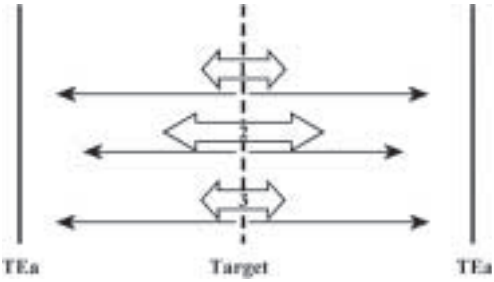


Figure 5-20. Z-Bars© illustrate method performance and ΔSEc for the WBC control results in Table 5-5.

illustrate method performance and critical systematic error. The TEa limits are set at 15% (per CLIA guidelines).

In this hypothetical example the error rate is categorized as high. This indicates that the process is frequently prone to significant errors that will cause the method to exceed quality specifications—and therefore we will need to hold back patient results. We have selected the QIK-QC Strategy Table© #3 (Figure 5-23) for N = 2 and a high error rate. Selecting the appropriate row from the QIK QC Strategy Table©, we choose a 1-2s single rule or a combination of 1-2s, R-4s, 4-1s, and 8x for each level of control. This is the maximum error detection we can obtain with QC rules.

In order to increase the probability of error detection, we examine the QC chart daily and increase control frequency. We may choose to run more controls in each run, or to increase the frequency of controls to every batch, for example, ev-

QIK-QC Strategy Table© #4 Error Rate = Low N = 3, 6, 9		
SEc	Single Rule	Multi-Rule
> 3.0	1-3.5s	1-3s/3-1s(w)
2.0 - 3.0	1-3.0s	1-2s/2-2s/R-4s/3-1s(w)
1.0 - 2.0	1-2.5s	1-3s/2-2s/R-4s/3-1s
< 1.0	1-2.0s	1-2s/R-4s/3-1s/12s

= Examine QC chart daily
= Increase control frequency
= Initiate corrective action

Figure 5-21. QIK-QC Strategy Table© #4 recommends appropriate QC rules and strategy for a method with a low error rate and controls of 3, 6, or 9.

Table 5-6. Prothrombin Time Control Results

Prothrombin Time	Control 1	Control 2
Mean	13.5	36.5
SD	0.50	1.50
Target	13.0	35.0
TEa	1.95	5.25
ΔSEc	1.3	0.9

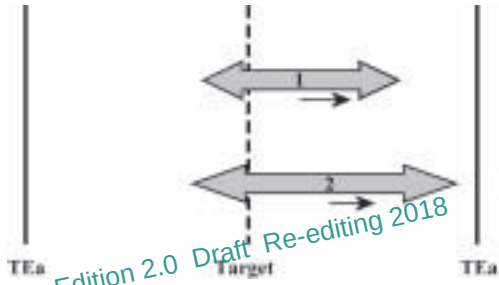


Figure 5-22. Z-Bars© illustrate method performance and ΔSEc for the prothrombin time control results in Table 5-6.

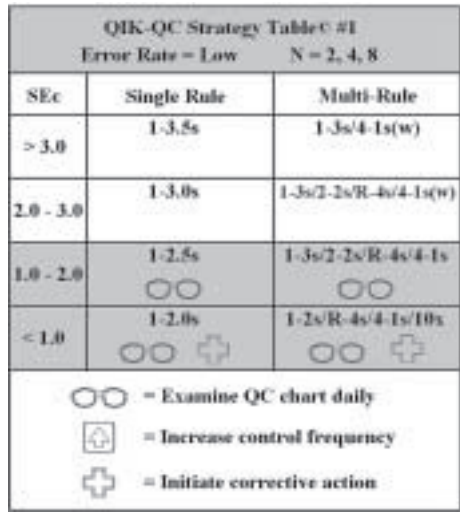


Figure 5-23. QIK-QC Strategy Table© #3 recommends appropriate QC rules and strategy for a method with a high error rate and controls of 2, 4, or 8.

ery two hours. As this method has a high error rate and low ΔSEc, it is important to initiate corrective action to improve method accuracy and precision and to reduce the overall error rate.

HIDDEN HAZARDS

The likelihood that a QC flag is a true reject or false reject. In an ideal world, every QC flag would signify a true reject; if no significant errors occur, we would see no QC flags. The probability of error detection and false rejection varies with the magnitude of change in the mean and SD and with the QC rules we elect to use. Table 5-7 illustrates how the probability of error detection and false rejection varies with ΔSEc and error rate in a hypothetical laboratory that tests 1,000 control samples for each analytical method each year. In this example, we have assumed that

- a method that experiences significant problems twice a year would be categorized as having a low error rate and would generate 4 QC results from "bad" runs with N = 2, and 6 "bad" results with N = 3;
- a method that experiences significant problems every two months would be categorized as having a moderate error rate and would generate 12 QC results from "bad" runs with N = 2, and 18 "bad" results with N = 3;
- a method that experiences significant problems every month would be categorized as having a high error rate and would generate 24 QC results from "bad" runs with N = 2, and 36 "bad" results with N = 3;

In Table 5-7:

- Examples are provided for error rates varying from low to high, as discussed above.
- Examples include ΔSEc of 4.5, 2.5, and 1.5 to simulate different critical shifts in the mean.

Table 5-7. Probability of error detection and false rejection varies with ΔSEc and error rate (N = 2).

Error rate	ΔSEc	QC Rule	Ped	Pfr	True Rejects	Detected in first run	False rejects
Low	4.5	3.5	100%	0%	4	4	0
Mod		3.5	100%	0%	12	12	0
High		3.0	100%	0%	24	24	0
Low	2.5	3.0	55%	0%	4	2	0
Mod		2.5	73%	3%	12	9	25
High		2.0	88%	10%	24	21	100
Low	1.5	2.5	28%	3%	4	1	25
Mod		2.0	50%	10%	12	6	100
High		2.0	50%	10%	24	12	100

- QC rules are selected from the QIK-QC Strategy Tables©.
- Ped and Pfr are estimated from power function graphs created by James Westgard.
- The number of true rejects is based on the error rate (2, 6, 12 times each year) multiplied by the number of controls per run.
- The number of errors detected in the first run is the number of true rejects multiplied by the Ped.
- The number of false rejects is the Pfr times the number of controls (1,000) tested each year.

When ΔSEc is 4.5, regardless of the error rate, we can use a 1-3.5s rule and be confident we will detect any significant errors in the first QC run, with no false positive QC flags. When ΔSEc is only 2.5, we select a 1-3s, 1-2.5s, or 1-2s rule as the error rate changes from low, to moderate, to high. Our confidence that we will detect a significant error in the first QC run is lower than with the ΔSEc of 4.5. Remember that if the mean shifts 4.5 SD, almost all points will be outside ± 3 SD and we will quickly see QC flags. When the mean shifts 2.5 SD, fewer points will fall outside the ± 2 SD or ± 3 SD limits. When we choose to use a 1-2.5s or 1-2s rule, the number of false positive QC flags increases. With two controls in a run, 10% of "good" runs will fall outside ± 2 SD, and 3% will exceed ± 2.5 SD. When the ΔSEc is only 2.5, and the method error rate is moderate or high, the number of false positive QC flags is higher than the number of true rejects. Methods with moderate accuracy and precision relative to quality specifications have a lower probability of detecting errors in the first analytical run, should a significant error occur.

When ΔSEc is only 1.5, we select a 1-2.5s or 1-2s rule to increase the probability of error detection as the error rate changes from low, to moderate, to high. Even with a 1-2s rule, our confidence that we will detect a significant error in the first QC run is lower than with the ΔSEc of 2.5. As we discussed above, with two controls in a run, 10% of "good" runs will fall outside ± 2 SD, and 3% will exceed ± 2.5 SD. When the ΔSEc is only 1.5, and the method error rate is moderate or high, the number of false positive QC flags is much higher than the number of true rejects. Methods with moderate accuracy and precision relative to quality specifications have a lower probability of detecting errors in the first analytical run, should a significant error occur. For these methods, QIK-QC Strategy Tables© recommend increasing control frequency, increasing visual examination of QC charts, and initiating corrective measures to improve method performance.

With three controls per run, as shown in Table 5-8, the number of true rejects increases to 6, 18, and 36 for low, moderate, and high error rates. The number of false positive QC flags also increases. As with two controls per run, we see that probability of error detection is higher when ΔSEc is higher.

Figure 5-24 shows variations in the number of QC flags with varying ΔSEc and error rates for two and three controls per run based on

- the number of "bad" runs or significant errors based on a low, moderate, or high error rate (The error rate of a method is independent of ΔSEc. A method

Table 5-8. Probability of error detection and false rejection varies with ΔSEc and error rate (N = 3).

Error rate	ΔSEc	Rule	Ped	Pfr	True Rejects	Detected in first run	False rejects
Low	4.5	3.5	100%	0%	6	6	0
Mod		3.5	100%	0%	18	18	0
High		3.0	100%	0%	36	36	0
Low	2.5	3.0	60%	0%	6	4	0
Mod		2.5	78%	3%	18	14	30
High		2.0	92%	15%	36	33	150
Low	1.5	2.5	35%	3%	6	2	30
Mod		2.0	60%	15%	18	11	150
High		2.0	60%	15%	36	22	150

may have the necessary accuracy and precision to easily meet quality specifications, yet be prone to frequent sources of error and have a high error rate.);

- the number of errors detected in the first run, based on the Ped of the QC rules selected (as the error rate increases or ΔSEc decreases, we select rules with a higher Ped); and
- the number of false rejects based on the Pfr of the QC rules selected (QC rules with a higher Ped also have a higher Pfr).

Notice that when ΔSEc is high (4.5), we detect all the "bad" results in the first run and see no false rejects with either two or three controls per run. Perhaps we should reconsider why we use three controls for some tests. This practice often dates back to older (possibly manual) methodology when we needed a third control

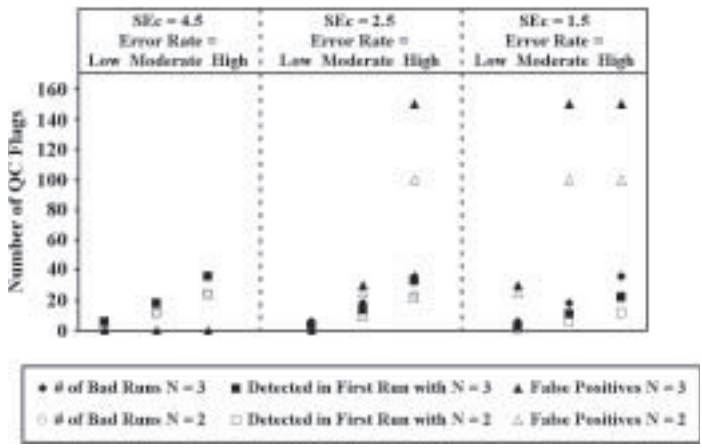


Figure 5-24. Probability of error detection and false rejection with varying ΔSEc and error rates.

to ensure performance in non-linear methods. By applying Performance-Driven Quality Control, we may be able to reduce the number of controls we analyze for some tests.

The number of "bad" results is constant. As ΔSEc decreases, the probability of detecting a significant error in the first run decreases. As ΔSEc decreases, we choose QC rules with a higher Ped and higher Pfr to increase the likelihood of detecting errors in the first QC run. Unfortunately, when Ped increases, Pfr also increases, and it is not uncommon to see more false rejects than true rejects when ΔSEc is low.

Looking objectively at the probability of error detection may provide incentive to:

1. use a QC process that makes us aware of ΔSEc values;
2. investigate ways to increase ΔSEc; and
3. improve error detection and reduce false rejection by selecting appropriate QC rules.

Can we use the same QC rule(s) all the time for each test? Despite the knowledge that analytical processes frequently experience changes in method performance, we often use the same QC rule(s) at all times. Figure 5-25 shows variations in the mean, SD, and ΔSEc of a single sodium control tested on the same analytical instrument over six months. When ΔSEc is low, as it was in June 1999, we can increase error detection by choosing QC rules with a high Ped, and examining QC charts more frequently. We can also initiate corrective action to reduce method bias and improve precision. When method performance improves, as it did in July 1999, we can choose QC rules with a lower Pfr and examine the QC chart less often. We can then re-direct these efforts to other methods that may require more attention at this time.

Can we use the same QC rule(s) for several instruments or laboratories? Different analytical processes or instruments are subject to independent factors

Date	*Mean*	*SD*	*SEc*	*Z-Bar*
15/09/99	142	1.07	2.75	
15/08/99	142.1	0.87	3.83	
15/07/99	143.2	0.68	4.32	
15/06/99	141	2.02	0.18	
15/05/99	140.9	1.4	0.91	
15/04/99	141.5	0.87	3.17	

Figure 5-25. Variations in mean, SD, and ΔSEc over six months for a sodium control with target of 142.3 mmol/L and TEa of 5.0 mmol/L.

Lab	Mean	SD	SEc	Z-Bar
1	2.95	0.036	3.88	
2	2.88	0.031	2.61	
3	2.81	0.027	0.39	
4	2.96	0.024	6.43	
5	2.90	0.020	5.98	
6	2.90	0.029	3.56	

Figure 5-26. Variations in mean, SD, and ΔSEc over six months for a calcium control with target of 2.95 mmol/L and TEa of 0.20 mmol/L.

that change method performance, yet we often decide to use the same QC rule(s) for several instruments in one or more laboratories.

Figure 5-26 shows variations in the mean, SD, and ΔSEc of a single calcium control tested on the same analytical instrument model in six laboratories. In laboratory #3, when ΔSEc is low, we can increase error detection by choosing QC rules with a high Ped and examining QC charts more frequently. We can also initiate corrective action to reduce method bias and improve precision. When ΔSEc is >3.0, as it is in laboratory #1, #4, #5, and #6, we can choose QC rules with a lower Pfr and examine the QC chart less often. By applying Performance-Driven Quality Control, we can "right-size" our QC efforts. The QC process becomes fluid and responds in real time to changes in the analytical process.

What are the costs associated with using an arbitrary 1-2s rule? In the study described above, of 2,711 controls results (mean and SD compared to target and TEa) monitored in 6 laboratories, 2,167 had an ΔSEc >3.0 and, with a low error rate, could use a 1-3.5s QC rule. Figure 5-27 shows that only 93 tests required a

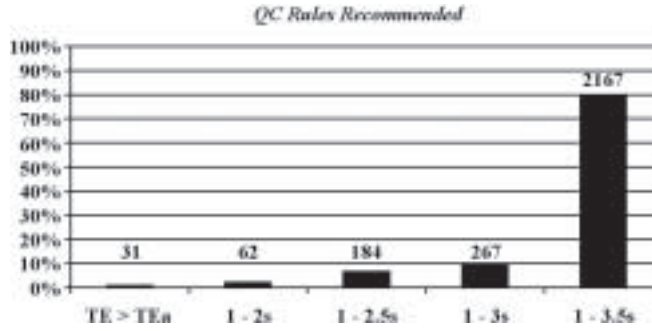


Figure 5-27. QC rules for 23 routine chemistry tests in six laboratories selected using Performance-Driven QC.

1-2s rule and performance improvement. Use of a 1-2.5s, 1-3s, and 1-3.5s QC rule reduces false positive QC flags.

In this study, 2,618 controls did not require a 1-2s rule. At only 30 QC runs per month, these controls represent 78,540 individual QC data points. A 5% false positive rate from the 1-2s rule would have produced 3,927 flags. At an estimated cost of only $10 for the labor and materials for each unnecessary repeat run or QC investigation, the cost of using a 1-2s rule when it is not required represents $39,270.

SUMMARY

When QC is arbitrary it does not assess method performance relative to goals and may not alert us to significant changes in the analytical process. Deliberate QC design ensures that we select QC processes to meet defined quality specifications.

A QC process should be designed to match the accuracy, precision, and error rate of each analytical system. This means selecting the number of controls, QC rules, and strategy to balance the changing ΔSEc and error rate of each method. Different QC rules will detect varying sizes and different types of error. Error detection is improved by increasing the number of controls tested (N). A Performance-Driven Quality Control system enables us to change QC rules, number of controls, control frequency, and other QC strategies to match current analytical performance and error rate.

Critical System Error (ΔSEc) is used to indicate method performance relative to a defined quality specification. A high ΔSEc indicates a method that can tolerate a large shift in the mean and requires a relatively lenient QC strategy. A low ΔSEc indicates a method that can tolerate only a small shift in the mean and requires a relatively stringent QC strategy.

As the overall error rate of a method increases, the need for QC rules and additional processes to detect error also increases. Q.I.K. QC Strategy Tables© provide recommendations for QC rules, control frequency (or N), frequency of examining QC charts, and the need for corrective action. These tables can help design QC processes that maximize error detection and minimize false rejection.