Implementation Guidance of Statistical Quality Control Rules for Quantitative Assays

Table of Contents

1.0 Scope and background

1.1 Scope

Quality Control (QC) aims to detect and correct deficiencies observed in routine laboratory processes which provides robust and reliable test results that will be used as part of delivering the best patient care possible. QC rules help to detect shifts, drifts or imprecisions in the QC measurement process.

This guidance supports laboratories in the selection and implementation of Quality Control (QC) rules specifically to those who do not have access to QC software which supports the selection of suitable QC rules.

Please note that this manual is not a comprehensive guidance on all aspects of QC. In addition, a laboratory has to ensure that its QC procedures fulfill local and global regulations (e.g. Clinical and Laboratory Standards Institute Guideline C24 [1] or ISO 15189 [2]).

1.2 Quality control requirements for quantitative assays

For quantitative assays specific to those co-developed with pharma partners, the assigned target ranges must be met. Depending on the assay, there are two or three levels of controls.

For certain assays used in patient management (for example, assays informing on a patient's eligibility for dosing of a drug), the user also needs to ensure that the laboratory-specific systematic bias, the laboratory-specific intermediate precision and the allowable total error [TEa] is within specification. These specifications are described in the 'Quality control' section of the assay instructions for use.

Note: this document, Implementation Guidance of Statistical Quality Control Rules for Quantitative Assays, is only relevant for those assays that make specific reference to it in their method sheet. Please refer to the 'Quality control' section of the method sheet to see if this document is relevant for the assay.

1.3 Structure of the guidance

The following sections outline the recommended experimental setup and evaluation of the initial laboratory-specific analytical performance, and the subsequent QC rules selection for routine QC monitoring, based on the Threshold Sigma Metric concept. Sections 2.0 and 3.0 describe the QC methodology when there are two respectively three control levels.

Section 2.0 applies only to assays with two control levels, while Section 3.0 applies only to assays with three control levels.

2.0 QC methodology for quantitative assays with 2 control levels

An assessment of the laboratory-specific analytical performance may be necessary prior to implementing an appropriate QC rule in a laboratory. For assays with 2 control levels, it is recommended to follow the procedure that is explained step-by-step in the following sections.

2.1 QC performance

2.1.1 QC experiment

The purpose of the QC experiment is to collect a robust data set of assay measurements by using both Levels of Controls (LCs), that is LC1 and LC2 to estimate the initial performance of the analytical system in terms of relative bias and imprecision. The initial performance has to fulfill certain QC requirements, in order to establish suitable QC rules, as described in the following sections.

1. It is preferred to collect at least $n = 20$ LC1 and at least $n = 20$ LC2 measurement results. For instruments that have more than one measuring entity, 20 LC1 results and 20 LC2 results should be collected per measuring entity.
2. It is preferred to distribute the LC1 and LC2 measurements over at least ten days with two runs per day. Since calibrations could add some source of variation it is preferred to conduct at least three calibrations within the duration of the experiment. If relevant, known additional sources of variation influencing the routine measurement process should be included in the experiment.

In the following example, a QC experiment in a laboratory that plans to conduct measurements with an assay for a hypothetical Biomarker "X" is illustrated and sample calculations are shown. In this example, it is assumed that the TEa for measurements of Biomarker "X" is given as 25%, with an allowable relative bias of ±12% and an allowable random error of 8%.

The laboratory performs the QC experiment using the both LCs with exemplary target values (TVs) (see Table 1).

	Target values [unit]
LC1	7.42
LC2	39.8

Table 1: Example target values for the two levels of the two LCs of an assay for Biomarker "X" used for patient management.

Each LC is measured on 10 days with two runs per day to generate 20 data points per measuring unit. Calibrations are performed on days 1, 4 and 7. The example data set of one measuring entity are shown in Table 2.

		LC1 [unit]	LC2 [unit]
Day 1	Run 1	8.03	40.9
	Run 2	7.82	41.3
Day 2	Run 1	7.94	40.4
	Run 2	7.54	42.8
Day 3	Run 1	7.52	41.2
	Run 2	7.54	41.5
Day 4	Run 1	7.91	40.9
	Run 2	8.29	42.9
Day 5	Run 1	7.44	42.6
	Run 2	7.46	41.4
Day 6	Run 1	7.56	40.2
	Run 2	7.44	40.4
Day 7	Run 1	7.68	42.6
	Run 2	7.89	38.5
Day 8	Run 1	7.83	42.2
	Run 2	7.59	41.0
Day 9	Run 1	7.73	42.6
	Run 2	7.89	43.2
Day 10	Run 1	7.61	40.7
	Run 2	7.95	42.8

Table 2: Data for the two LCs over 10 days with two runs per day.

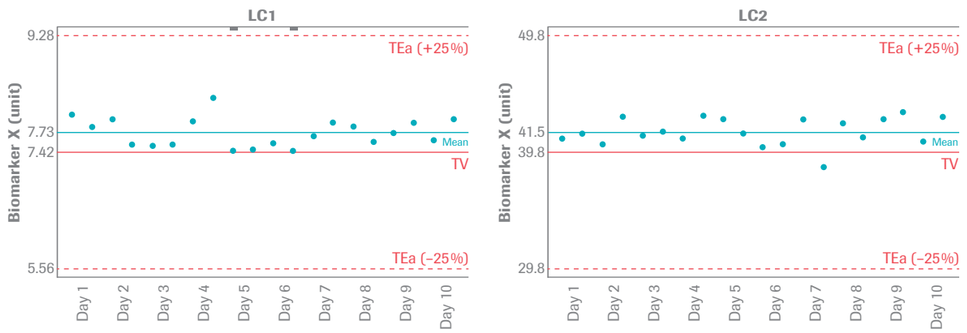

Figure 1: Levey-Jennings charts for the data of the QC experiment for each LC. The red solid lines show the target values (TV) for the corresponding LC and the dashed red lines show the lower and upper TEa limits of ±25%. The blue line shows the laboratory-specific mean value for each LC.

It is recommended to plot the data over days and runs (Levey-Jennings chart) for each LC as shown in Figure 1.

In this example, it appears that the LC measurement results are close to their target values with small dispersion and are also safely within the TEa limits of ±25%.

2.1.2 QC performance quantification

This section provides information on how to evaluate analytical performance of a laboratory based on the results of the QC experiment (shown in Figure 1) and on how to quantify the 'distance' between the TEa limits and the analytical performance. This quantified distance is commonly known as Sigma metric and denoted by σ. Note that the Sigma metric σ is not only used to quantify the QC performance of a laboratory, but also used to select the QC rules as shown in Section 2.2. The Sigma metric σ for the data observed in the QC experiment is calculated using the following steps[1].

Note for customers of cobas® analyzers containing at least two measuring entities and using them in routine: As mentioned previously, the full QC experiment should be performed on each measuring entity. It is then possible to select a separate QC rule for each measuring entity or to select a pooled QC rule for both measuring entities. If monitoring is planned to be performed using a pooled QC rule, the LC measurement results of the QC experiment from both measuring entities need to be pooled for each LC. For the calculation of the Sigma metric σ shown in the following, an adjustment of the number n of assay LC measurement results is then necessary. For example, in case of two measuring entities, $n = 2 \times 20 = 40$ LC measurement results per LC are available.

1) Please note that for some calculations, rounded interim results are reported and used in subsequent calculations. Usage of unrounded values may lead to slightly different values.

Step 1: Calculate the laboratory specific mean and standard deviation for each LC

The mean is calculated as the sum of the LC measurement results divided by the number n of LC measurement results. The calculated mean for a LC is formally expressed as:

$$Mean = \frac{1}{n} \times (x_1 + x_2 + \ldots + x_{n-1} + x_n) = \frac{1}{n} \times \sum_{i=1}^{n} x_i$$

where x_i denotes a valid measurement of a LC and n the total number of valid measurements of a LC. In the example data set, the calculated mean of each LC is shown in Table 3.

The standard deviation (SD) quantifies the variation of the LC measurements around the calculated mean. To calculate the standard deviation, take the sum of the squared differences between the LC measurements and the calculated mean. This sum is divided by the number of LC measurements per LC minus one. Finally, taking the square root of this fraction gives the standard deviation. The calculation of the standard deviation is formally expressed as:

$$SD = \sqrt{\frac{1}{n\text{-}1} \times \sum_{i=1}^{n} (x_i - Mean)^2}$$

In the example data set, the calculation of the SD is shown in Table 4.

		LC1 [unit]	LC2 [unit]
Day 1	Run 1	8.03	40.9
	Run 2	7.82	41.3
Day 2	Run 1	7.94	40.4
	Run 2	7.54	42.8
Day 3	Run 1	7.52	41.2
	Run 2	7.54	41.5
Day 4	Run 1	7.91	40.9
	Run 2	8.29	42.9
Day 5	Run 1	7.44	42.6
	Run 2	7.46	41.4
Day 6	Run 1	7.56	40.2
	Run 2	7.44	40.4
Day 7	Run 1	7.68	42.6
	Run 2	7.89	38.5
Day 8	Run 1	7.83	42.2
	Run 2	7.59	41.0
Day 9	Run 1	7.73	42.6
	Run 2	7.89	43.2
Day 10	Run 1	7.61	40.7
	Run 2	7.95	42.8
Sum of LC results		154.66	830.1
Number n of LC results		20	20
Mean		$\frac{154.66}{20} \approx 7.73$	$\frac{830.1}{20} \approx 41.5$

Table 3: Calculation of the mean for each LC of the data obtained from the QC experiment shown in Table 2.

		LC1 [unit]	Squared differences LC1	LC2 [unit]	Squared differences LC2
Day 1	Run 1	8.03	$(8.03 - 7.73)^2$	40.9	$(40.9 - 41.5)^2$
	Run 2	7.82	$(7.82 - 7.73)^2$	41.3	$(41.3 - 41.5)^2$
Day 2	Run 1	7.94	$(7.94 - 7.73)^2$	40.4	$(40.4 - 41.5)^2$
	Run 2	7.54	$(7.54 - 7.73)^2$	42.8	$(42.8 - 41.5)^2$
Day 3	Run 1	7.52	$(7.52 - 7.73)^2$	41.2	$(41.2 - 41.5)^2$
	Run 2	7.54	$(7.54 - 7.73)^2$	41.5	$(41.5 - 41.5)^2$
Day 4	Run 1	7.91	$(7.91 - 7.73)^2$	40.9	$(40.9 - 41.5)^2$
	Run 2	8.29	$(8.29 - 7.73)^2$	42.9	$(42.9 - 41.5)^2$
Day 5	Run 1	7.44	$(7.44 - 7.73)^2$	42.6	$(42.6 - 41.5)^2$
	Run 2	7.46	$(7.46 - 7.73)^2$	41.4	$(41.4 - 41.5)^2$
Day 6	Run 1	7.56	$(7.56 - 7.73)^2$	40.2	$(40.2 - 41.5)^2$
	Run 2	7.44	$(7.44 - 7.73)^2$	40.4	$(40.4 - 41.5)^2$
Day 7	Run 1	7.68	$(7.68 - 7.73)^2$	42.6	$(42.6 - 41.5)^2$
	Run 2	7.89	$(7.89 - 7.73)^2$	38.5	$(38.5 - 41.5)^2$
Day 8	Run 1	7.83	$(7.83 - 7.73)^2$	42.2	$(42.2 - 41.5)^2$
	Run 2	7.59	$(7.59 - 7.73)^2$	41.0	$(41.0 - 41.5)^2$
Day 9	Run 1	7.73	$(7.73 - 7.73)^2$	42.6	$(42.6 - 41.5)^2$
	Run 2	7.89	$(7.89 - 7.73)^2$	43.2	$(43.2 - 41.5)^2$
Day 10	Run 1	7.61	$(7.61 - 7.73)^2$	40.7	$(40.7 - 41.5)^2$
	Run 2	7.95	$(7.95 - 7.73)^2$	42.8	$(42.8 - 41.5)^2$
Sum of squared differences			≈1.0206		≈27.21
Number n –1 of LC results			$20 - 1 = 19$		$20 - 1 = 19$
SD			$\sqrt{\frac{1.0206}{19}} \approx 0.23$		$\sqrt{\frac{27.21}{19}} \approx 1.20$

Table 4: Calculation of the standard deviation (SD) for each LC of the data obtained from the QC experiment shown in Table 2.

Step 2: Calculate bias, coefficient of variation, and Sigma metric σ for each LC and determine smallest Sigma metric

For the calculation of the Sigma metric σ, both the bias and observed coefficient of variation (CV) for each LC are needed. These quantities are calculated using the means and standard deviations for each LC.

The bias is calculated for each LC by

$$Bias\,[\%] = \left| \frac{Mean - TV}{TV} \right| \times 100$$

where TV denotes the Roche target value of the corresponding LC and |.| denotes the absolute value[2].

2) The observed relative bias would typically be compared against an established international standard or reference method, but might not exist for every assay. The peer group mean from external quality assessment (EQA) schemes may be used to assess the bias but currently the availability of EQA schemes for any assay might be limited.

The coefficient of variation is calculated for each LC by

$$CV\,[\%] = \frac{SD}{Mean} \times 100.$$

If both, Bias[%] is less than or equal to 12% and CV[%] is less than or equal to 8%, then the QC requirements are met.

Based on Bias[%] and CV[%], the Sigma metric σ is calculated for each LC using the following formula:

$$\sigma = \frac{TEa\,[\%] - Bias\,[\%]}{CV\,[\%]}$$

Subsequently, the smaller one of the two Sigma metrics obtained for the two LCs is selected and used to derive the QC rules. Alternatively, it is possible to use the worst Bias[%] across the two LCs and the worst CV[%] across the two LCs and calculate the Sigma metric σ accordingly.

In the example data set, the Bias[%], CV[%], Sigma metric σ for each LC, and the smallest Sigma metric across the two LCs are calculated as follows:

	LC1	LC2
Target Value (TV) [unit]	7.42	39.8
Mean [unit]	7.73	41.5
Standard Deviation (SD) [unit]	0.23	1.20
Bias[%]	$\left\lvert \frac{7.73 - 7.42}{7.42} \right\rvert \times 100 = 4.18\%$	$\left\lvert \frac{41.5 - 39.8}{39.8} \right\rvert \times 100 = 4.27\%$
Bias[%] less than 12%?	**Yes**	**Yes**
CV[%]	$\frac{0.23}{7.73} \times 100 = 2.98\%$	$\frac{1.20}{41.5} \times 100 = 2.89\%$
CV[%] less than 8%?	**Yes**	**Yes**
Sigma Metric σ	$\frac{25\% - 4.18\%}{2.98\%} = 6.99$	$\frac{25\% - 4.27\%}{2.89\%} = 7.17$
Smallest Sigma Metric σ	Minimum (6.99, 7.17) = **6.99**	

Table 5: Calculation of the bias and coefficient of variation (CV) and verification of QC requirements (TEa criteria). Calculation of the Sigma metric for each LC and determination of the smallest Sigma metric σ.

In the example data set, all QC requirements are met and the Sigma metric σ for each LC can be calculated. If one of the Biases[%] is greater than 12% or one of the two CVs[%] is greater than 8%, troubleshooting and root-cause analysis to determine and eliminate the sources of bias or variation has to be conducted (see section 4.0). Troubleshooting and root-cause analysis also might be required if the smallest Sigma metric value is less than 4.36, since a laboratory will not be able to identify a feasible QC rule under this condition (see section 4.0).

2.2 QC rule selection

The application of one or more suitable QC rules is preferred for routine measurements with quantitative Roche assays where results are used for patient management.

The following section describes the procedure to select a QC rule. In general, a laboratory can select a QC rule from a set of rules based on the laboratory's initial performance. The performance is quantified by the Sigma metric σ.

A set of QC rules widely used in laboratory routine are the so-called 'Westgard rules' proposed by Westgard et al.[3] Other types of QC rules exist and can be used, as long as they fulfill the criteria explained below. This guidance document focuses primarily on within-run Westgard rules, in particular for the manual selection of QC rules based on the Threshold Sigma Metric (TSM) explained later. These within-run Westgard rules immediately detect significant changes in the measurement conditions.

A QC rule applied to quantitative assays for patient management should meet the two following probabilistic criteria:

- The QC rule should detect a gradual or sudden change in the QC measurement process leading to a violation of the QC requirements with a sufficiently large probability (usually 90% or more) in a QC run. This probability is known as probability of error detection (P_{ed}) or power. The QC requirements are violated if the TEa limits are exceeded (i.e., if the probability of a measurement outside the limits is equal to or above 5%).
- A QC rule should have a sufficiently small probability (usually 5% or less) of falsely rejecting a QC run when the QC measurement process is in-control. This probability is known as probability of false rejection (P_{fr}).

Note for customers who have access to a QC rule selection software: A laboratory could use a QC rule selection software based on the results and probabilistic criteria as specified above. Customers who have a QC rule selection software can skip the following paragraphs and continue with Section 2.3.

The following section describes how an appropriate QC rule can be selected if the laboratory has no access to a QC rule selection software.

Every Westgard rule has a fixed probability of false rejection (P_{fr}), which is below 5% for all QC rules discussed in this section. Each of these rules is associated with a so-called Threshold Sigma Metric (TSM). If the smallest Sigma metric σ from the QC experiment is greater than or equal to the TSM of a QC rule, the 90% power criterion is fulfilled and the QC rule can be used for QC monitoring.

Table 6 provides a list of common QC rules with corresponding TSM and P_{fr}.

QC rule	TSM	P_{fr}
1-4s	6.12	0.01%
1-3.5s	5.62	0.09%
1-3s	5.12	0.54%
1-2.81s	4.93	0.99%
1-2.5s	4.62	2.47%
1-2.24s	4.36	4.96%
1-3s | 2-2s	4.77	0.63%

Table 6: Set of Westgard rules with corresponding Threshold Sigma Metric (TSM) and probability of false rejection (P_{fr}).

3) For more information on Westgard rules, and the theory behind them, see references [3] and [4] in the references or https://www.westgard.com/westgard-rules.htm

QC rules such as type 1-*k*s are commonly used ones which reject a QC run if the QC measurement result of at least one LC is more than k laboratory-specific SDs ($k \in \{2.24, 2.5, 2.81, 3, 3.5, 4\}$) above or below its laboratory-specific mean. The 2-2s rule rejects a QC run if the QC measurement results of both LCs are at least within +2 SDs or –2 SDs from its lab-specific means. The 1-3s | 2-2s QC rule rejects a QC run if the 1-3s or 2-2s QC rule is triggered. If more than one QC rule from Table 6 is suitable, the rule with the lowest P_{fr} should be selected.

In the example data set, the smallest Sigma metric σ is compared with the TSM values and Table 6 is amended accordingly:

QC rule	TSM	Is smallest Sigma metric σ = 6.99 ≥ TSM?	P_{fr}
1-4s	6.12	**Yes**	0.01%
1-3.5s	5.62	**Yes**	0.09%
1-3s	5.12	**Yes**	0.54%
1-2.81s	4.93	**Yes**	0.99%
1-2.5s	4.62	**Yes**	2.47%
1-2.24s	4.36	**Yes**	4.96%
1-3s | 2-2s	4.77	**Yes**	0.63%

Table 7: Benchmarking the smallest Sigma metric σ of 6.99 observed in the example data set with the Threshold Sigma Metric (TSM).

Table 7 shows that the smallest Sigma metric σ of 6.99 is larger than the TSM for all QC rules. In this example, a laboratory would choose the 1-4s rule, because it is a single rule with the lowest P_{fr}.

2.3 QC rule implementation

The previously selected QC rule from section 2.2 is used to calculate the laboratory specific control ranges. The calculation of the lower and upper limits for each LC are based on the factor k associated with the QC rule and are calculated as follows:

- *Lower LC limit = Mean − k × SD*
- *Upper LC limit = Mean + k × SD*

Note that in the case of the 2-2s QC rule, the factor k is 2. Therefore, for the 1-3s | 2-2s QC rule, two lower and upper limits have to be considered.

In the example data set, the limits for each LC are calculated using the selected 1-4s QC rule as shown in Table 8.

LC	Mean [unit]	SD [unit]	Factor k	Lower LC limit [unit]	Upper LC limit [unit]
LC1	7.73	0.23	4	7.73 − 4 × 0.23 = 6.81	7.73 + 4 × 0.23 = 8.65
LC2	41.5	1.20	4	41.5 − 4 × 1.20 = 36.7	41.5 + 4 × 1.20 = 46.3

Table 8: Calculation of the lower and upper limits of the 1-4s rule for both LCs.

These ranges for each LC can be entered into the **cobas**® analyzer software. Note that only 1-*k*s rules can be entered on the **cobas**® analyzer. QC rules that are more complex can be used with the lab software.

Note for customers of cobas e 411 analyzer:
Only 1-3s rule can be entered into the analyzer. In order to convert a QC rule of the type 1-*k*s into the required 1-3s rule, use the following equation:

$$SD^a = \frac{k}{3} \times \mathrm{SD}$$

The converted SD, SD^a, can be entered into the analyzer. In our data example, which uses the 1-4s QC rule, the converted SDs for the two LCs are calculated by

$$\frac{4}{3} \times 0.23 \approx 0.31 \qquad \frac{4}{3} \times 1.20 \approx 1.60$$

By applying this conversion, the resulting QC limits are identical to the QC limits of table 8 (disregarding rounding differences).

2.4 Routine use of the assay

After following the above guidance, the Roche quantitative assays are ready for routine and daily use in a patient management setting. Perform routine monitoring and corrective actions, if needed, according to laboratory protocols.

3.0 QC methodology for quantitative assays with 3 control levels

An assessment of the laboratory-specific analytical performance may be necessary prior to implementing an appropriate QC rule in a laboratory. It is recommended to follow the procedure that is explained step-by-step in the following sections.

3.1 QC performance

3.1.1 QC experiments

The purpose of the QC experiment is to collect a robust data set of assay measurements by using three Levels of Controls (LCs), that is LC1, LC2 and LC3 to estimate the initial analytical system performance based on relative bias and imprecision. The initial performance has to fulfill certain QC requirements, in order to establish suitable QC rules, as described in the following sections.

1. It is preferred to collect at least $n = 20$ LC1, at least $n = 20$ LC2 and at least $n = 20$ LC3 measurement results. For instruments that have more than one measuring entity, 20 LC1, 20 LC2, and 20 LC3 results should be collected per measuring entity.
2. It is preferred to distribute the LC1, LC2 and LC3 measurements over at least ten days with two runs per day. Since calibrations could add some source of variation it is preferred to conduct at least three calibrations within the duration of the experiment. If relevant, known additional sources of variation influencing the routine measurement process should be included in the experiment.

In the following example, a QC experiment in a laboratory that plans to conduct measurements with an assay for a hypothetical Biomarker “X” is illustrated and sample calculations are shown. In this example, it is assumed that the TEa for measurements of Biomarker “X” is given as 25%, with an allowable relative bias of ±12% and an allowable random error of 8%.

The laboratory performs the QC experiment using the three LCs with exemplary target values (TVs) (see Table 9).

	Target values [unit]
LC1	7.42
LC2	39.8
LC3	60.3

Table 9: Example target values for the three levels of the LCs of an assay for Biomarker “X” used for patient management.

Each LC is measured on 10 days with two runs per day to generate 20 data points per measuring unit. Calibrations are performed on days 1, 4 and 7. The example data of one measuring entity are shown in Table 10.

		LC1 [unit]	LC2 [unit]	LC3 [unit]
Day 1	Run 1	8.18	41.0	65.3
	Run 2	7.93	41.5	66.5
Day 2	Run 1	8.07	40.5	66.1
	Run 2	7.58	43.3	62.9
Day 3	Run 1	7.56	41.4	65.2
	Run 2	7.58	41.8	63.4
Day 4	Run 1	8.04	41.1	63.4
	Run 2	8.50	43.4	66.6
Day 5	Run 1	7.47	43.0	63.2
	Run 2	7.49	41.6	65.4
Day 6	Run 1	7.61	40.3	62.7
	Run 2	7.47	40.5	63.3
Day 7	Run 1	7.75	42.9	64.3
	Run 2	8.01	38.3	60.2
Day 8	Run 1	7.94	42.6	63.8
	Run 2	7.65	41.1	64.1
Day 9	Run 1	7.82	43.0	63.2
	Run 2	8.01	43.7	64.3
Day 10	Run 1	7.67	40.8	65.0
	Run 2	8.08	43.2	59.3

Table 10: Data for the three LCs over 10 days with two runs per day.

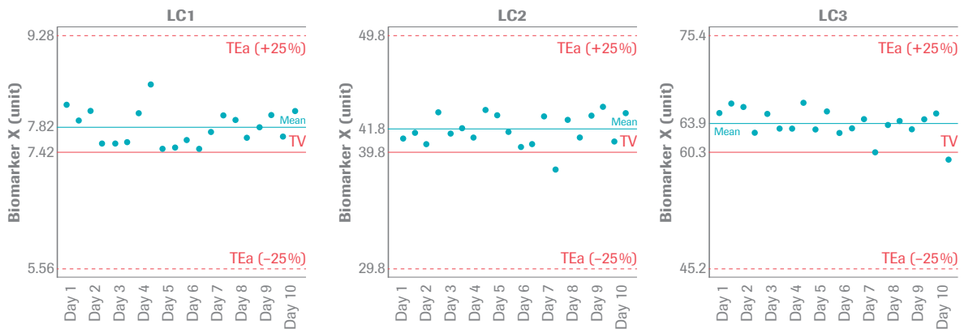

Figure 2: Levey-Jennings charts for the data of the QC experiment for each LC. The red solid lines show the target values (TV) for the corresponding LC and the dashed red lines show the lower and upper TEa limits of ±25%. The blue line shows the laboratory-specific mean value for each LC.

It is recommended to plot the data over days and runs (Levey-Jennings chart) for each LC as shown in Figure 2.

In this example, it appears that the LC measurement results are close to their target values with little dispersion. In addition, it shows that the LC results are safely within the TEa limits of ±25%.

3.1.2 QC performance quantification

This section provides information on how to evaluate analytical performance of a laboratory based on the QC experiment results (shown in Figure 2) and on how to quantify the 'distance' between the TEa limits and the analytical performance. This quantified distance is commonly known as Sigma metric and denoted by σ. Note that the Sigma metric σ is not only used to quantify the QC performance of a laboratory, but also used to select the QC rules as shown in Section 3.2. The Sigma metric σ for the data observed in the QC experiment is calculated using the following steps[4].

Note for customers of cobas® analyzers containing at least two measuring cells and using them in routine: As mentioned above, the full QC experiment should be performed on each measuring cell. It is possible to select a separate QC rule for each measuring cell or to select a pooled QC rule for both measuring cells. If monitoring is planned to be performed using a pooled QC rule, the LC measurement results of the QC experiment from both measuring cells need to be pooled for each LC. For the calculation of the Sigma metric σ shown in the following, an adjustment of the number n of assay LC measurement results is then necessary. For example, in case of two measuring cells, $n = 2 \times 20 = 40$ LC measurement results per LC are available.

4) Please note that for some calculations, rounded interim results are reported and used in subsequent calculations. Usage of unrounded values may lead to slightly different values.

Step 1: Calculate the laboratory specific mean and standard deviation for each LC

The mean is calculated as the sum of the LC measurement results divided by the number n of LC measurement results. The calculated mean for a LC is formally expressed as:

$$Mean = \frac{1}{n} \times (x_1 + x_2 + \ldots + x_{n-1} + x_n) = \frac{1}{n} \times \sum_{i=1}^{n} x_i$$

where x_i denotes a valid measurement of a LC and n the total number of valid measurements of a LC. The calculated mean of each LC for the example data is shown in Table 11.

The standard deviation (SD) quantifies the variation of the LC measurements around the calculated mean. To calculate the standard deviation, take the sum of the squared differences between the LC measurements and the calculated mean. This sum is divided by the number of LC measurements per LC minus one. Finally, taking the square root of this fraction gives the standard deviation. The calculation of the standard deviation is formally expressed as:

$$SD = \sqrt{\frac{1}{n\text{-}1} \times \sum_{i=1}^{n} (x_i - Mean)^2}$$

For the example data, the calculation of the SD is shown in Table 12.

		LC1 [unit]	LC2 [unit]	LC3 [unit]
Day 1	Run 1	8.18	41.0	65.3
	Run 2	7.93	41.5	66.5
Day 2	Run 1	8.07	40.5	66.1
	Run 2	7.58	43.3	62.9
Day 3	Run 1	7.56	41.4	65.2
	Run 2	7.58	41.8	63.4
Day 4	Run 1	8.04	41.1	63.4
	Run 2	8.50	43.4	66.6
Day 5	Run 1	7.47	43.0	63.2
	Run 2	7.49	41.6	65.4
Day 6	Run 1	7.61	40.3	62.7
	Run 2	7.47	40.5	63.3
Day 7	Run 1	7.75	42.9	64.3
	Run 2	8.01	38.3	60.2
Day 8	Run 1	7.94	42.6	63.8
	Run 2	7.65	41.1	64.1
Day 9	Run 1	7.82	43.0	63.2
	Run 2	8.01	43.7	64.3
Day 10	Run 1	7.67	40.8	65.0
	Run 2	8.08	43.2	59.3
Sum of LC results		156.4	835.0	1,278.2
Number n of LC results		20	20	20
Mean		$\frac{156.4}{20} \approx 7.82$	$\frac{835.0}{20} \approx 41.8$	$\frac{1,278.2}{20} \approx 63.9$

Table 11: Calculation of the mean for each LC of the data obtained from the QC experiment shown in Table 10.

		LC1 [unit]	Squared diff. LC1	LC2 [unit]	Squared diff. LC2	LC3 [unit]	Squared diff. LC3
Day 1	Run 1	8.18	$(8.18 - 7.82)^2$	41.0	$(41.0 - 41.8)^2$	65.3	$(65.3 - 63.9)^2$
	Run 2	7.93	$(7.93 - 7.82)^2$	41.5	$(41.5 - 41.8)^2$	66.5	$(66.5 - 63.9)^2$
Day 2	Run 1	8.07	$(8.07 - 7.82)^2$	40.5	$(40.5 - 41.8)^2$	66.1	$(66.1 - 63.9)^2$
	Run 2	7.58	$(7.58 - 7.82)^2$	43.3	$(43.3 - 41.8)^2$	62.9	$(62.9 - 63.9)^2$
Day 3	Run 1	7.56	$(7.56 - 7.82)^2$	41.4	$(41.4 - 41.8)^2$	65.2	$(65.2 - 63.9)^2$
	Run 2	7.58	$(7.58 - 7.82)^2$	41.8	$(41.8 - 41.8)^2$	63.4	$(63.4 - 63.9)^2$
Day 4	Run 1	8.04	$(8.04 - 7.82)^2$	41.1	$(41.1 - 41.8)^2$	63.4	$(63.4 - 63.9)^2$
	Run 2	8.50	$(8.50 - 7.82)^2$	43.4	$(43.4 - 41.8)^2$	66.6	$(66.6 - 63.9)^2$
Day 5	Run 1	7.47	$(7.47 - 7.82)^2$	43.0	$(43.0 - 41.8)^2$	63.2	$(63.2 - 63.9)^2$
	Run 2	7.49	$(7.49 - 7.82)^2$	41.6	$(41.6 - 41.8)^2$	65.4	$(65.4 - 63.9)^2$
Day 6	Run 1	7.61	$(7.61 - 7.82)^2$	40.3	$(40.3 - 41.8)^2$	62.7	$(62.7 - 63.9)^2$
	Run 2	7.47	$(7.47 - 7.82)^2$	40.5	$(40.5 - 41.8)^2$	63.3	$(63.3 - 63.9)^2$
Day 7	Run 1	7.75	$(7.75 - 7.82)^2$	42.9	$(42.9 - 41.8)^2$	64.3	$(64.3 - 63.9)^2$
	Run 2	8.01	$(8.01 - 7.82)^2$	38.3	$(38.3 - 41.8)^2$	60.2	$(60.2 - 63.9)^2$
Day 8	Run 1	7.94	$(7.94 - 7.82)^2$	42.6	$(42.6 - 41.8)^2$	63.8	$(63.8 - 63.9)^2$
	Run 2	7.65	$(7.65 - 7.82)^2$	41.1	$(41.1 - 41.8)^2$	64.1	$(64.1 - 63.9)^2$
Day 9	Run 1	7.82	$(7.82 - 7.82)^2$	43.0	$(43.0 - 41.8)^2$	63.2	$(63.2 - 63.9)^2$
	Run 2	8.01	$(8.01 - 7.82)^2$	43.7	$(43.7 - 41.8)^2$	64.3	$(64.3 - 63.9)^2$
Day 10	Run 1	7.67	$(7.67 - 7.82)^2$	40.8	$(40.8 - 41.8)^2$	65.0	$(65.0 - 63.9)^2$
	Run 2	8.08	$(8.08 - 7.82)^2$	43.2	$(43.2 - 41.8)^2$	59.3	$(59.3 - 63.9)^2$
Sum of squared differences			≈ 1.506		≈ 35.90		≈ 65.50
Number n –1 of LC results			$20 - 1 = 19$		$20 - 1 = 19$		$20 - 1 = 19$
SD			$\sqrt{\frac{1.506}{19}} \approx 0.282$		$\sqrt{\frac{35.90}{19}} \approx 1.37$		$\sqrt{\frac{65.50}{19}} \approx 1.86$

Table 12: Calculation of the standard deviation (SD) for each LC of the data obtained from the QC experiment shown in Table 10.

Step 2: Calculate bias, coefficient of variation, and Sigma metric σ for each LC and determine smallest Sigma metric

To calculate the Sigma metric σ, both the bias and observed coefficient of variation (CV) for each LC are needed. These quantities are calculated based on the means and standard deviations for each LC.

The bias is calculated for each LC by

$$Bias\,[\%] = \left| \frac{Mean - TV}{TV} \right| \times 100$$

where TV denotes the Roche target value of the corresponding LC and |.| denotes the absolute value[5].

5) The observed relative bias would typically be compared against an established international standard or reference method, but might not exist for every assay. The peer group mean from external quality assessment (EQA) schemes may be used to assess the bias but currently the availability of EQA schemes for any assay might be limited.

The coefficient of variation is calculated for each LC using the following equation:

$$CV\,[\%] = \frac{SD}{Mean} \times 100.$$

If both, Bias[%] is less than or equal to 12% and CV[%] is less than or equal to 8%, then the QC requirements are met.

Based on Bias[%] and CV[%], the Sigma metric σ is calculated for each LC using the following formula:

$$\sigma = \frac{TEa\,[\%] - Bias\,[\%]}{CV\,[\%]}$$

	LC1	LC2	LC3
Target Value (TV) [unit]	7.42	39.8	60.3
Mean [unit]	7.82	41.8	63.9
Standard Deviation (SD) [unit]	0.282	1.37	1.86
Bias[%]	$\left\lvert \frac{7.82 - 7.42}{7.42} \right\rvert \times 100 = 5.39\%$	$\left\lvert \frac{41.8 - 39.8}{39.8} \right\rvert \times 100 = 5.03\%$	$\left\lvert \frac{63.9 - 60.3}{60.3} \right\rvert \times 100 = 5.97\%$
Bias[%] less than 12%?	**Yes**	**Yes**	**Yes**
CV[%]	$\frac{0.282}{7.82} \times 100 = 3.61\%$	$\frac{1.37}{41.8} \times 100 = 3.28\%$	$\frac{1.86}{63.9} \times 100 = 2.91\%$
CV[%] less than 8%?	**Yes**	**Yes**	**Yes**
Sigma Metric σ	$\frac{25\% - 5.39\%}{3.61\%} = 5.43$	$\frac{25\% - 5.03\%}{3.28\%} = 6.09$	$\frac{25\% - 5.97\%}{2.91\%} = 6.54$
Smallest Sigma Metric σ	Minimum (5.43, 6.09, 6.54) = **5.43**		

Table 13: Calculation of the bias and coefficient of variation (CV) and verification of QC requirements (TEa criteria). Calculation of the Sigma metric for each LC and determination of the smallest Sigma metric σ.

Subsequently, the smallest one of the three Sigma metrics obtained for the three LCs is selected and used to derive the QC rules. Alternatively, it is possible to use the worst Bias[%] across the three LCs and the worst CV[%] across the three LCs and to calculate the Sigma metric σ accordingly.

In the example data set, the Bias[%], CV[%], Sigma metric σ for each LC and the smallest Sigma metric across the three LCs are calculated as follows:
In the example data set, all QC requirements are met and the Sigma metric σ for each LC can be calculated. If one of the Biases[%] is greater than 12% or one of the three CVs[%] is greater than 8%, troubleshooting and root-cause analysis to determine and eliminate the sources of bias or variation has to be conducted (see section 4.0). Troubleshooting and root-cause analysis also may be required if the smallest Sigma metric value is less than 4.12, since a laboratory will not be able to identify a feasible QC rule under this condition (see section 4.0).

3.2 QC rule selection

The application of one or more suitable QC rules is preferred for routine measurements with Roche quantitative assays where results are used for patient management.

The following section describes the procedure to select a QC rule. In general, a laboratory can select a QC rule from a set of rules based on the initial laboratory performance. The performance is quantified by the Sigma metric σ.

A set of QC rules widely used in laboratory routine are the so-called 'Westgard rules' proposed by Westgard et al.[6] Other types of QC rules exist and can be used for enhanced QC, as long as they fulfill the criteria explained below. However, this guidance document focuses primarily on within-run Westgard rules, in particular for the manual selection of QC rules based on the Threshold Sigma Metric (TSM) explained further in this section. These within-run Westgard rules immediately detect significant changes in the measurement conditions.

A QC rule applied to quantitative assays for patient management should meet the two following probabilistic criteria:

- The QC rule should detect a gradual or sudden change in the QC measurement process leading to a violation of the QC requirements with a sufficiently large probability (usually 90% or more) in a QC run. This probability is known as probability of error detection (P_{ed}) or power. The QC requirements are violated if the TEa limits are exceeded (i.e., if the probability of a measurement outside the limits is equal to or above 5%).
- A QC rule should have a sufficiently small probability (usually 5% or less) of falsely rejecting a QC run when the QC measurement process is in-control. This probability is known as probability of false rejection (P_{fr}).

Note for customers who have access to a QC rule selection software: A laboratory could use a QC rule selection software based on the results and probabilistic criteria as specified above. Customers who have a QC rule selection software can skip the following paragraphs and continue with Section 3.3.

The following section describes how an appropriate QC rule can be selected if the laboratory has no access to a QC rule selection software.

Every Westgard rule has a fixed probability of false rejection (P_{fr}), which is below 5% for all QC rules discussed in this section. Each of these rules are associated with a so-called Threshold Sigma Metric (TSM). If the smallest Sigma metric σ from the QC experiment is greater than or equal to the TSM of a QC rule, the 90% power criterion is fulfilled and the QC rule can be used for QC monitoring.

In Table 14, a list of common QC rules with corresponding TSM and P_{fr} is provided.

QC rule	TSM	P_{fr}
1-4s	5.73	0.02%
1-3.5s	5.23	0.14%
1-3s	4.73	0.81%
1-2.5s	4.23	3.68%
1-2.39s	4.12	4.97%
1-3s | 2/3-2s	4.27	1.08%

Table 14: Set of Westgard rules with corresponding Threshold Sigma Metric (TSM) and probability of false rejection (P_{fr}).

6) For more information on Westgard rules, and the theory behind them, see references [3] and [4] in the references or https://www.westgard.com/westgard-rules.htm

QC rules such as type 1-*k*s are commonly used ones which reject a QC run if the QC measurement result of at least one LC is more than k laboratory-specific SDs (k ∈ {2.39, 2.5, 3, 3.5, 4}) above or below its laboratory-specific mean. The 2/3-2s rule rejects a QC run if the QC measurement results of at least two out of the three LCs are at least within +2 SDs or -2 SDs from its lab-specific means. The 1-3s | 2/3-2s QC rule rejects a QC run if the 1-3s or 2/3-2s QC rule is triggered. If more than one QC rule from Table 14 is suitable, the rule that provides the lowest P_{fr} should be selected.

In the example data set, the smallest Sigma metric σ is compared with the TSM values and Table 14 is amended accordingly:

QC rule	TSM	Is smallest Sigma metric σ = 5.43 ≥ TSM?	P_{fr}
1-4s	5.73	**No**	0.02%
1-3.5s	5.23	**Yes**	0.14%
1-3s	4.73	**Yes**	0.81%
1-2.5s	4.23	**Yes**	3.68%
1-2.39s	4.12	**Yes**	4.97%
1-3s | 2/3-2s	4.27	**Yes**	1.08%

Table 15: Benchmarking the smallest Sigma metric σ of 5.43 observed in the data example with the Threshold Sigma Metric (TSM).

Table 15 shows that the smallest Sigma metric σ of 5.43 is larger than the TSM for all QC rules except the 1-4s rule. In this example, a laboratory would choose the 1-3.5s rule, because it is a single rule that provides the lowest P_{fr}.

3.3 QC rule implementation
The previously selected QC rule from section 3.2 is used to calculate the laboratory specific control ranges. The calculation of the lower and upper limits for each LC are based on the factor k associated with the QC rule and are calculated as follows:

- *Lower LC limit = Mean – k × SD*
- *Upper LC limit = Mean + k × SD*

Note that in the case of the 2/3-2s QC rule, the factor k is 2. Thus, for the 1-3s | 2/3-2s QC rule, two lower and upper limits have to be considered.

In the example data set, the limits for each LC are calculated using the selected 1-3.5s QC rule as shown in Table 16.

LC	Mean [unit]	SD [unit]	Factor k	Lower LC limit [unit]	Upper LC limit [unit]
LC1	7.82	0.28	3.5	7.82 – 3.5 × 0.28 = 6.84	7.82 + 3.5 × 0.28 = 8.80
LC2	41.8	1.37	3.5	41.8 – 3.5 × 1.37 = 37.0	41.8 + 3.5 × 1.37= 46.6
LC3	63.9	1.86	3.5	63.9 – 3.5 × 1.86 = 57.4	63.9 + 3.5 × 1.86 = 70.4

Table 16: Calculation of the lower and upper limits of the 1-3.5s rule for all three LCs.

These ranges for each LC can be entered into the **cobas**® analyzer software. Note that only 1-ks rules can be entered on the **cobas**® analyzer. QC rules that are more complex can be used with the lab software.

Note for customers of cobas e 411 analyzer:
Only 1-3s rule can be entered into the analyzer. In order to convert a QC rule of the type 1-ks into the required 1-3s rule, use the following equation:

$$SD^a = \frac{k}{3} \times \text{SD}$$

The converted SD, SD^a, can be entered into the analyzer. In our data example, which uses the 1-3.5s QC rule, the converted SDs for the two LCs are calculated by

$$\frac{3.5}{3} \times 0.28 \approx 0.33 \qquad \frac{3.5}{3} \times 1.37 \approx 1.60$$

$$\frac{3.5}{3} \times 1.86 \approx 2.17$$

By applying this conversion, the resulting QC limits are identical to the QC limits of table 16 (disregarding rounding differences).

3.4 Routine use of the assay
After following the above guidance, the Roche quantitative assays are ready for routine and daily use in a patient management setting. Perform routine monitoring and corrective actions, if needed, according to laboratory protocols.

4.0 Troubleshooting

If the assay performance does not meet the minimum QC requirements, an investigation must be performed on potential sources of imprecision or bias. Potential error sources of imprecision may include, but are not limited to:

1. Operational errors
   a. Handling of reagent, controls, calibrators and working solutions
   b. Instrument performance should be checked. Maintenance should be performed according to schedule
2. Environmental conditions (e.g. temperature, humidity, storage conditions, reagent storage conditions)

Potential sources of bias are usually systematic error sources and may include, but are not limited to:

1. Failure to recalibrate, according to the user instructions, if the lot number for the reagent or LCs has changed
2. Use of expired reagents
3. An aged measuring cell demonstrating diminishing performance
4. Wrong handling of QC material

If all attempts to meet the QC requirements are unsuccessful, call Roche customer service.

5.0 Change of reagent lot or control lot

If the reagent lot or control lot is changed, it is recommended that the laboratory continues to analyze samples according to the existing QC rules. Once the laboratory has collected 20 independent QC measurements for each LC, the Sigma metric should be recalculated. If the quality requirements identified are not met, troubleshooting should be performed (see section 4.0) or a new QC experiment should be performed.

References

1 *CLSI (2016). Statistical Quality Control for Quantitative Measurement Procedures: Principles and Definitions. 4th ed. CLSI guideline C24. Wayne, P.A.: Clinical and Laboratory Standards Institute.*

2 *ISO (2012). Medical laboratories – Requirements for quality and competence. ISO 15189:2012. Geneva: International Organization for Standardization.*

3 *Westgard, J., Barry, P., Hunt, M., Groth, T. (1981). A multi-rule Shewhart chart for quality control in clinical chemistry.* ***Clinical Chemistry 27,*** *493-501.*

4 *Westgard, J., Groth, T., Aronsson, T., Falk, H., Verdier, dC. (1977). Performance characteristics of rules for internal quality control: probabilities for false rejection and error detection.* ***Clinical Chemistry 23,*** *1857-1867.*

5 *Schoenmakers, C.H.H., Naus, A.J.M., Vermeer, H.J., Loon, D., Steen, G. (2011). Practical application of Sigma Metrics QC procedures in clinical chemistry.* ***Clinical Chemistry in Laboratory Medicine 49,*** *1837-1843.*

Published by:
Roche Diagnostics International Ltd
CH-6343 Rotkreuz
Switzerland

diagnostics.roche.com/cobas