

Test Group Summary Report
C37 Colour in Water
October 2019 PT Round

1.0 THE PROFICIENCY TESTING REPORT

The Proficiency Testing Report consists of two parts:

CALA Proficiency Testing Report: This report contains participant-specific evaluations and other confidential information. This report is emailed to participants at the end of the study.

Test Group Summary Report: A Test Group Summary Report is created for each test group at the end of the study. These reports contain more detailed information on the study than are found in the participant-specific CALA Proficiency Testing Report. These reports do not contain any confidential information and are made available on the CALA web site.

2.0 DEFINITIONS

The participant-specific CALA Proficiency Testing Report contains some terms that new participants may not be familiar with.

Code:	The registration code that is unique to each analyte that a participant is registered for.
App:	If a participant is accredited by CALA, this three digit number is the appendix number that the accredited method is assigned to.
N:	The number of participants results that were used to calculate the summary statistics. This excludes qualified data (e.g., <) and any results that were flagged as outliers.
Assigned:	The Assigned Value is the robust mean of the reported results, outliers excluded. This is often referred to as the “target” value.
± u:	The uncertainty of the assigned value.
Reported:	The result reported by the participant.
s:	The Standard Deviation of Proficiency Assessment. This value is used to determine the acceptance limits for the PT evaluation.
z-Score:	A value assigned to each reported result that is a measure of how much it deviates from the Assigned Value.
Score:	The composite score of the four results reported for each analyte. It is normalized to a score out of 100.
Bias:	A flag assigned if bias is detected using the re-scaled z-score procedure.

3.0 SCORING SYSTEM

Participant performance is evaluated for each proficiency testing sample by a quantitative method that is consistent with ISO/IEC 17043:2010 *Conformity assessment – General requirements for proficiency testing*, the *International Harmonized Protocol for Proficiency Testing of (Chemical) Analytical Laboratories* (2006), and ISO 13528:2015 *Statistical methods for use in proficiency testing by interlaboratory comparisons*.

The following is a brief description of the evaluation procedure used by CALA. The detailed evaluation procedure is described in PT15-03 - *CALA PT Program - Procedures*, which is available on the CALA website.

3.1 Homogeneity and Stability Assessment

Homogeneity and stability are assessed using participant data. Regression analysis is performed on reported result against order of sample production (Homogeneity) and reported result against date of analysis (Stability). If the slope is significantly different than zero for either then the Standard Deviation of Proficiency Assessment (s) is increased until the impact is minimized.

3.2 The z score

A "z-score" is calculated for each reported result as follows:

$$z = \frac{(x - \bar{X})}{s}$$

where: x = participant's result;

\bar{X} = the Assigned Value for the sample;

s = the Standard Deviation for Proficiency Assessment.

The assigned value \bar{X} is generally estimated from the inter-laboratory Robust mean after outliers due to obvious gross errors (e.g., reported in wrong units) have been removed.

The Standard Deviation for Proficiency Assessment, s , is determined as follows:

1. The inter-laboratory Robust standard deviation (stdev) is calculated using reported results, obvious outliers removed;
2. The expected inter-laboratory standard deviation ($s!$) is estimated from regression equations derived from previous studies (see PT15-05-CALA PT Program – *Regression Equations* for details);
3. If $s!$ is higher than stdev then $s!$ is used in the z score equation;
4. If $s!$ is lower than stdev then stdev is used in the z score equation;
5. When a laboratory reports its detection limit, s will be estimated using a pooled variance procedure that uses both the inter-laboratory data and the reported detection limit.

3.2 Composite (PT) Score

Since each PT study involves four or two separate samples of distinct concentrations for each test, it is necessary to calculate a composite PT score for each test to determine overall performance. The composite score is calculated by first averaging the absolute z-scores for the four results and then calculating a final score as $100 + (-15 \times \text{avg } |z|)$.

Acceptable PT Scores equal or exceed 70.

3.3 Identifying Bias

The proficiency testing report provides flags for bias. These are determined using the re-scaled z-score procedure.

$$RSZ = \frac{\sum z}{\sqrt{N}}$$

where z = the z- score
 N = the number of samples

Flags are assigned for each test group/parameter combination as follows:

$RSZ \geq -2$ and ≤ 2	no flag assigned
$RSZ > 2$	H (High)
$RSZ > 3$	VH (Very High)
$RSZ < -2$	L (LOW)
$RSZ < -3$	VL (Very Low)

3.4 Deviations from Evaluation Procedure

Other than changes to the Standard Deviation of Proficiency Assessment due to homogeneity or stability flags, any deviation from the published evaluation procedure is described on the cover page(s) of the final PT report.

4.0 STUDY SPECIFIC DATA SUMMARY

The following pages provide more detailed information about the study indicated in the cover page of this report than is found in the participant-specific CALA Proficiency Testing Report. The graphical representations and the statistical summaries by method are based upon the data after outliers have been removed.

4.1 Overall Data Distribution

Kernel density plots are generated for each data set. These plots are a statistical way to represent the overall data distribution and are used to visualize deviations from normality and bi-modality.

4.2 z-Score Plots

The z-scores for each sample are ranked in increasing order and plotted. When the data is normally distributed, the plot should show a slight sigmoidal curve, with an equal number of points above zero as below.

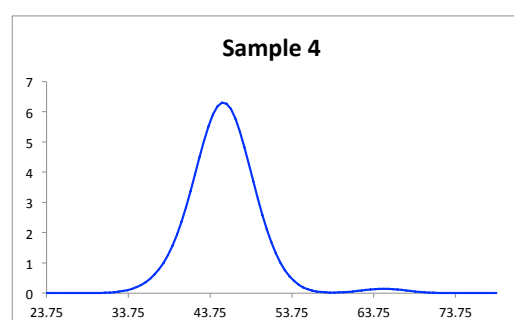
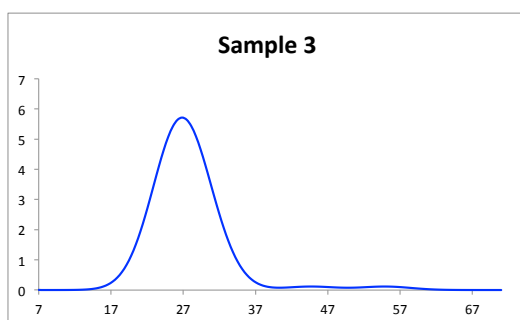
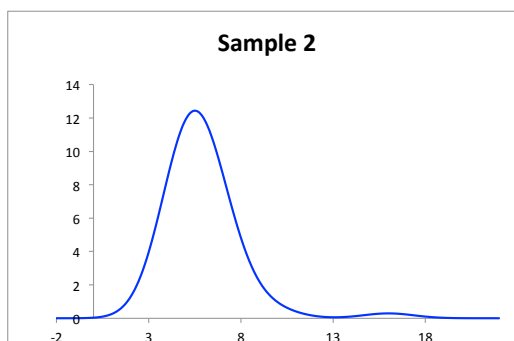
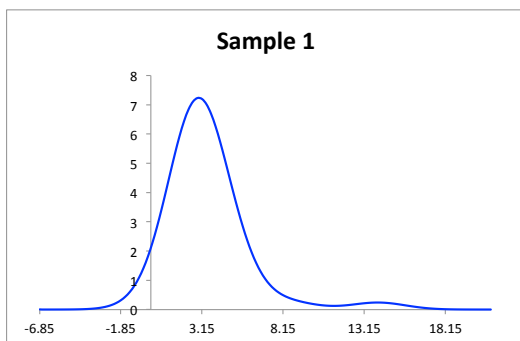
4.3 Statistical Summary by Method

Statistical summaries are provided for the four most commonly reported methods. The statistics included in these tables for mean and standard deviation are the mathematical values, not the robust values. As such, there will be slight differences between these and those displayed in the participant specific CALA Proficiency Testing Report.

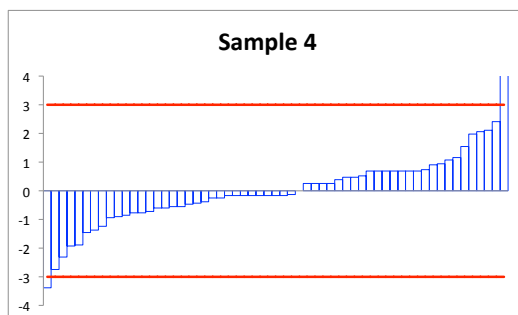
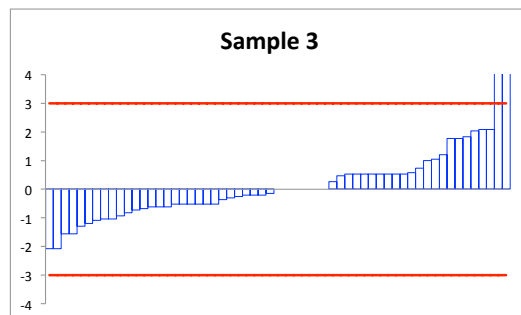
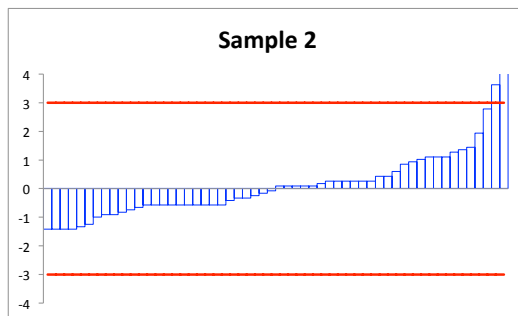
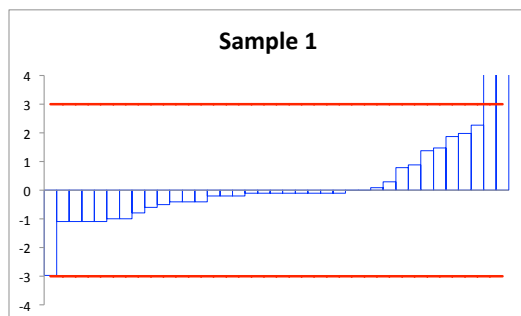
In addition to this, z-Score plots are provided for the four most common methods. As these plots use the Assigned Value and Standard Deviation for Proficiency Assessment estimated using all of the data, any method that does not have approximately the same number of positive and negative z-scores is an indication that the method may be biased relative to other methods.

True Colour

Kernel Density Plots



z-Score Plots



Summary Statistics by Four Most Common Methods

Sample 1					
Method	All*	COLOR	VC	CLR-A	
Median	3.11	3	2.7	2.9	0
Stdev	1.01	2.33	1.57	0	0
Number	37	33	3	1	0
z > 3	2	2	0	0	0
z 2 - 3	2	2	0	0	0

Sample 2					
Method	All*	COLOR	VC	CLR-A	
Median	5.69	5.45	7	5.8	0
Stdev	1.19	1.9	2.05	0.987	0
Number	56	48	3	4	0
z > 3	2	2	0	0	0
z 2 - 3	1	0	1	0	0

Sample 3					
Method	All*	COLOR	VC	CLR-A	
Median	27	26.6	28	27.6	0
Stdev	1.92	5.02	1.73	2.63	0
Number	59	50	4	4	0
z > 3	2	2	0	0	0
z 2 - 3	5	4	1	0	0

Sample 4					
Method	All*	COLOR	VC	CLR-A	
Median	45.4	45.1	42.5	46.7	0
Stdev	2.33	3.67	3.75	1.32	0
Number	59	50	4	4	0
z > 3	2	1	1	0	0
z 2 - 3	5	4	1	0	0

* Values for the median and standard deviation are the robust mean and acceptance limit deviation. For method statistics, the arithmetic mean and standard deviation are used.

z-Score Plots by Method - Sample 4

