

PT15-03 – CALA PT Program - Procedures  
**Revision 1.14** – June 13, 2019



**CALA**  
Proficiency Testing

# TABLE OF CONTENTS

<b>1.0</b>	<b>INTRODUCTION .....</b>	<b>1</b>
<b>2.0</b>	<b>PROCEDURES.....</b>	<b>1</b>
2.1	Posting Of PT Schedule On Web Site .....	1
2.2	Notification Of PT Sample Shipment.....	1
2.3	Review Of PT Sample Instructions .....	2
2.4	Reference Laboratory Competence (PTM).....	2
2.5	Confirm Design Value .....	2
2.6	Review Of Web-Data-Entry Comments.....	2
2.7	Flag All Non-Detects And Non-Reports.....	2
2.8	Outlier Removal .....	3
2.9	Produce the Internal Summary Reports.....	3
2.10	Produce and Distribute Preliminary .....	4
2.11	Examination of Participant Data .....	4
2.11.1	Use of Participant Data to Assess Homogeneity and Stability .....	5
2.11.2	Examination of Data Distribution .....	6
2.11.3	Examination of Summary Statistics from Different Methods .....	6
2.11.4	Production and Review of Proficiency Testing Summary Reports (PTM) .	6
2.12	Authorize Release Of Reports .....	7
2.12.1	Produce PT Notice .....	7
2.12.2	PT Round Approval .....	7
2.13	Produce The Final PT Reports .....	7
2.13.1	Verification of non-automated data entries.....	7
2.13.2	Review of Report Format and Random Check of Participant reports .....	8
2.13.3	Issuing the Final Report .....	8
2.14	Review Of Corrective Action Reports .....	8
2.15	Failure To Provide Acceptable Car.....	8
2.16	Information Kept On Each PT Round .....	8
<b>3.0</b>	<b>CORRECTIVE/REMEDIAL ACTION FOR THE PT ROUND .....</b>	<b>9</b>
3.2	Preventive Actions.....	10
3.3	Correction Reports .....	10
<b>APPENDIX I:</b>	<b>PT EVALUATION PROCEDURE .....</b>	<b>11</b>
AI.1	The Assigned Value ( $\bar{X}$ ) .....	11
AI.2	The Inter-laboratory Standard Deviation (stdev).....	11
AI.3	The Regression Equation Standard Deviation (s!) .....	11
AI.4	The Assigned Deviation (s) .....	11
AI.5	z-Scores.....	11
AI.6	Calculate Composite PT Score.....	13
AI.7	Estimating and Flagging Bias .....	13
AI.8	Standard Uncertainty of the Assigned Value:.....	14
AI.9	Interpretation of PT Results .....	14
AI.10	Custom Reports .....	14
<b>APPENDIX II:</b>	<b>PT EVALUATION FOR MICROBIOLOGY (PRESENCE/ABSENCE) AND PCB AROCLORS.....</b>	<b>15</b>
All.1	Microbiology by presence/absence .....	15
All.2	PCB Aroclors.....	15

# CALA PT PROGRAM - PROCEDURES

## 1.0 INTRODUCTION

This document outlines the steps followed in the course of a PT round. The collaborators perform the functions as outlined in their respective contracts.

Most CALA documents referenced in this document are available in the CALA web-site library ([www.cala.ca/library.html](http://www.cala.ca/library.html)).

## 2.0 PROCEDURES

The following section details the chronologic steps in the conduct of a CALA PT round.

### 2.1 Posting Of PT Schedule On Web Site

By November of each year a draft schedule is produced for the following year. This is forwarded to the collaborators and affected CALA staff for review. Once approved, the schedule is posted to the CALA web-site ([www.cala.ca/pt\\_ship\\_schedule.html](http://www.cala.ca/pt_ship_schedule.html)). In general, the deadline for changes is four weeks prior to shipping, and the deadline for reporting is four to five weeks after shipping.

#### PT Sample Shipment Schedule\*

Test groups, parameters, matrices, units and concentrations can found in [Section 3.0 of our Program Document](#).

2010 PT Schedule					
	OCTOBER 2009	JANUARY	MARCH	JUNE	OCTOBER
Invoice/Ship Notice	August 6	November 10, 2009	January 7	April 15	August 12
Registration Deadline	September 18	December 18, 2009	February 12	May 21	September 17
Shipping	October 19	January 18	March 15	June 21	October 18
Reporting Deadline	November 20	February 19	April 16	July 23	November 19

### 2.2 Notification Of PT Sample Shipment

At least six weeks prior to a scheduled PT round, participants are notified by email of the pending shipment and provided with a link to their current registration list. This notice is provided with an invoice for the upcoming PT round.

## 2.3 Review Of PT Sample Instructions

At least four weeks prior to shipping, the PT sample instructions are reviewed for continued applicability. Any modified instruction sheets are uploaded to the web site ([www.cala.ca/Instructions.html](http://www.cala.ca/Instructions.html)). The only instructions that must be changed every round are the General PT Instructions and the toxicology instructions (C11, C12 and C13).

## 2.4 Reference Laboratory Competence (PTM)

The current scopes of accreditation are reviewed for each Reference laboratory used by Environment Canada to ensure that they are still competent. Copies of these scopes are maintained with each round file. This review may take place anytime throughout the round cycle.

## 2.5 Confirm Design Value

The design values are confirmed by the Collaborator with verification data from the Reference laboratory. This comparison is intended to only identify gross differences that may have resulted from a serious error in sample production. For test groups other than the microbiology PT, two randomly selected samples are analysed for every production lot.

## 2.6 Review Of Web-Data-Entry Comments

Following the reporting deadline, and prior to the evaluation of any participant data, the comments entered by participants during web-data-entry are reviewed and, where necessary, corrections are made to the raw data. Recurring comments may result in further investigation. This is also an opportunity for participants to add analytes to their PT registration.

When corrections are made, or when results are reported through the comments field of the web-entry page, these entries are verified for accuracy.

**Note (1):** The web-data-entry system is structured to limit the number of decimal places that are entered during the submission of PT results. If the maximum number of decimal places is exceeded, the lab cannot submit the data.

**Note (2):** A number ending in a five (5) is always rounded upwards in the CALA database.

## 2.7 Flag All Non-Detects And Non-Reports

All results that were reported with a <, > or that were not submitted are temporarily removed from the initial set of raw data for the purpose of calculating the performance statistics. To prevent biased estimates, all chemistry data that is reported as zero are treated as non-submitted results.

## 2.8 Outlier Removal

Steps i through v are performed and are solely for the purpose of the Preliminary Report. Outliers are flagged and temporarily removed using the Grubbs Test:

- i) all results for a unique test group/sample ID/analyte combination are sorted in ascending order ( $x_1, x_2, \dots, x_n$ );
- ii) the arithmetic mean ( $\bar{X}$ ) and standard deviation ( $s$ ) of the data set are calculated;
- iii)  $T$  is calculated for the first and last result in the sorted series as;

$$T = \frac{|(x_i - \bar{X})|}{s}$$

- iv) if the largest  $T$  value is greater than the value from the Grubbs Table, the point is temporarily removed;
- v) steps i) to iv) are repeated until no more points are removed.

## 2.9 Produce the Internal Summary Reports

The internal summary reports consist of:

- o *Reference Value Report*: a statistical summary with automated flags (see below for information included in this report);
- o *Proficiency Summary by Parameter Report*: provides a summary of z-scores above 2 and unacceptable PT scores for each analyte;
- o *Test Value Report*: tabulates all reported results for each analyte, sorted in increasing order; and,
- o *Tentative Participant Reports*: .pdf reports that are the same format as the final reports but only used for internal purposes. Statistical procedures used to produce this report are found in Appendix I.

The columns in the *Reference Value Report* are as follows:

**Number of Participants (N)**: The number of results used to calculate assigned value and standard deviation.

**Median**: the median (middle value) provided subsequent to removal of outlier data points. If N is even, the average of the middle two data points is used.

**Adjusted Mean ( $\bar{X}$ )**: the arithmetic mean calculated subsequent to removal of outlier data points.

**Design Value**: the concentration that was intended for the sample.

**Assigned Deviation Value (s)**: The inter-laboratory standard deviation (stdev) from the reported results (outliers removed) is estimated from,

**Warning Limit for Standard Deviation\***: the upper value for the inter-laboratory standard deviation; this value is determined from the regression equations estimated for the warning limits.

**Regression Equation Standard Deviation (s!)** is estimated using the regression equation estimated from historic studies (see PT15-05-CALA PT Program - Regression Equations).

**Std Dev Used:** PF if the regression equation value is used and C if the inter-laboratory standard deviation is used.

**Std Dev Flag\*:** Indicates when the inter-laboratory standard deviation exceeds the warning limit.

**Ref Val Flag\*:** Indicates when the design value differs significantly from the consensus mean.

**Outlier Count:** Indicates the number of results flagged as low outliers and high outliers.

\* These flags are no longer used for decision making as they have been made redundant with the use of robust statistics to produce the final evaluations.

These reports are examined to identify any obvious problems prior to issuing the preliminary reports. This involves checking:

- The *Reference Value Report*: Specific attention is given to a large number of outliers at either the high end or the low end, suggesting a possible skewing or bi-modality;
- The *Reference Value Report*: to identify any gross differences between the design value and the consensus mean.
- The *Proficiency Summary by Parameter Report*: for unusually high number of z-scores above 2 or an unusually large number of unacceptable PT scores (>15%).
- A selection of Tentative Reports to ensure that formatting is appropriate and that unacceptable scores seem appropriate. Specific attention will be given to test groups that have different evaluation procedures, such as PCBs and C05B;
- The *Test Value Report* for an unusually high number of non-detects.

If no obvious problems are identified, the PTM authorizes the release of the Preliminary PT Report. When obvious issues are identified they are investigated and addressed prior to issuing the *Preliminary Report*.

## 2.10 Produce and Distribute Preliminary

Following the procedures detailed in Appendix I, *Preliminary PT Reports* are prepared in csv format and emailed to participants. This is typically done within one week of the PT round reporting deadline and is done before a thorough examination of the data is completed.

## 2.11 Examination of Participant Data

Prior to approval of the *Final PT Reports*, participant data is examined for every analyte, in every sample, in every test-group. The *PT34\_Data\_Examination.xls* spreadsheet is used to calculate the robust statistics and assess homogeneity and stability. The *PT Test Group Reports.xls* spreadsheet is used to generate the Test Group Summary Reports.

As part of this examination, a *Proficiency Testing Summary Report* (PT35) is produced for each test-group. Each report contains summaries of the following examinations. Procedures for use of these spreadsheet are found in PT36-*Use of PT Data Examination Spreadsheets*.

### 2.11.1 Use of Participant Data to Assess Homogeneity and Stability

The uncertainty associated with sample homogeneity and sample stability should not contribute significantly to the overall uncertainty of the PT evaluation.

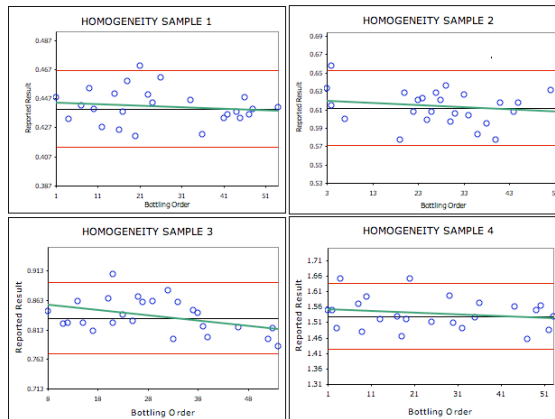
For each production lot of PT samples, regression analysis is conducted on participant reported result against bottling order. A possible homogeneity problem will display a trend in the data (increasing or decreasing) where the slope of the regression is significantly different than zero at  $\alpha = 0.05$  ( $p < 0.05$  when regression analysis is conducted) and the ratio of the maximum deviation from the assigned value to the standard deviation of proficiency testing is greater than 0.5.

As well, CALA assesses stability by plotting the reported result against the date of analysis using the same spreadsheet. This approach has the benefit of taking conditions during sample shipping into account.

A possible stability problem will display a trend in the data (increasing or decreasing) where the slope of the regression is significantly different than zero at  $\alpha = 0.05$  ( $p < 0.05$  when regression analysis is conducted) and the ratio of the maximum deviation from the assigned value to the standard deviation of proficiency testing is greater than 0.5.

When homogeneity or stability is flagged as a possible problem, the standard deviation of proficiency testing is increased to the point at which the ratio mentioned above is  $< 0.5$ . The picture below shows a spreadsheet screen capture for CO2B-3 Nickel that required a slight increase in the standard deviation to minimize the possible affect of inhomogeneity.

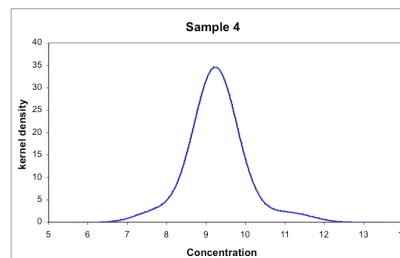
<b>HOMOGENEITY CONCLUSION</b>		Mean	Robust Mean	Median	% Diff
Sample 1	HOMOGENEOUS	0.440	0.4393	0.4393	0.16
Sample 2	HOMOGENEOUS	0.61	0.6088	0.608	0.65
Sample 3	NOT HOMOGENEOUS USE 0.042	0.83	0.829	0.8252	0.94
Sample 4	HOMOGENEOUS	1.53	1.528	1.526	0.26



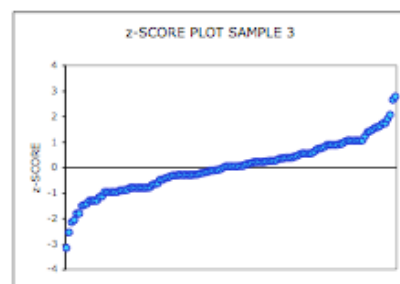
## 2.11.2 Examination of Data Distribution

One of the statistical assumptions of the CALA PT Scheme is that the data reported does not show bi-modality and approximates a normal distribution. This is examined through a combination of graphical and statistical means.

A kernel density plot is produced for each set of results. The only data removed when producing this plot are the obvious blunders (e.g., reporting in the wrong units). This type of plot is good for identifying modality and skewing problems.



As well, a ranked z-score plot is produced. If either of these shows significant skewing or, in the case of the z-score plot, a discontinuity, a comparison is made between the arithmetic mean, the median and the robust mean. For the estimates of the median and robust mean, only obvious gross errors are removed prior to calculation.



When participation levels are less than 11, or if there are a large number of non-detects, the data is examined to ensure that the assigned value is an acceptable estimate. If there is any doubt, the sample will not be evaluated.

## 2.11.3 Examination of Summary Statistics from Different Methods

Although PT evaluation uses all participant data, regardless of method, summary statistics are also produced and examined for the most commonly used methods. This table includes the overall arithmetic mean and standard deviation as well as the mean and standard deviation for up to four of the most common methods.

Summary Statistics by Four Most Common Methods					
Sample 1					
Method	All*	ENZYM-SUB	MF-DC	MF	MF-ENDO
Median	48.4	50	48	50	48
Stdev	9.2	10.4	7.9	6.27	9.71
Number	157	49	33	26	23
z  > 3	0	0	0	0	0
z  2 - 3	8	5	0	0	1

Whenever one method is found to be obviously and recurrently biased relative to other methods, laboratories using this method may be excluded from the evaluation and participants notified that the PT is not appropriate for that method. When this is done, a note is also included in P02-04 – CALA Program Description – PT Catalogue.

## 2.11.4 Production and Review of Proficiency Testing Summary Reports (PTM)

As part of the examination of data detailed above, a *Test Group Summary Report* (PT35) is produced for each test group. Each report contains:

- Summary of the evaluation procedure;
- Kernel Density plots;



- z-score plots;
- Statistical summary of the most commonly used methods; and,
- Z-score plots for the most commonly used analytical methods.

This examination is used to make decisions about the acceptability of the default evaluation procedure or if a modification is required. These are also of value to participants in investigating an unacceptable PT evaluation.

## 2.12 Authorize Release Of Reports

### 2.12.1 Produce PT Notice

If there are any deviations from the published PT evaluation procedure (e.g., excluding an analyte or a sample) or when there is other information that might be of value to the participant (e.g., a test group going from Pilot to Live status), a PT Notice is produced. This notice will provide details of this information and, where necessary, provide actions taken and rationale for these actions. This notice is appended to the cover page(s) of the confidential report.

### 2.12.2 PT Round Approval

After review of all data, the PTM authorizes the production and release of the final PT reports. This approval will detail all modifications to the normal evaluation procedure. These may include, but are not limited to:

- The use of an assigned value or assigned deviation that is different than consensus derived values;
- Removal of one or more samples from the evaluation procedure; or
- Removal of one or more analyte from the evaluation procedure.

The *Test Group Summary Reports* are posted to the CALA web-site for access by any interested party.

## 2.13 Produce The Final PT Reports

The participant specific Final PT Reports are produced using the evaluation procedure detailed in Appendix I.

### 2.13.1 Verification of non-automated data entries

Whenever participant results are reported in the comments field of the web-data-entry system, they are manually transcribed to an EXCEL template for uploading to the database for subsequent evaluation. These transcriptions are verified by comparing a database download of these entries against the print-out of the participants comments.

### **2.13.2 Review of Report Format and Random Check of Participant reports**

A random selection of participant reports is examined to ensure that formatting is correct and consistent. This will include page numbering, spacing, content and checking that any deviations are clearly indicated.

### **2.13.3 Issuing the Final Report**

Following the verification of any manual changes, the PTM authorizes the issuing of the Final Reports.

## **2.14 Review Of Corrective Action Reports**

CALA accredited participants are provided with blank corrective action report forms (CAR) for analytes that have unacceptable performance (see PT15-02-CALA PT Program - Policies for details). These are reviewed for completeness. In reviewing the corrective action reports, the a judgment is not normally made on the effectiveness of the corrective action.

Acceptability of the report is generally based on an indication that a thorough investigation has been made, an attempt to identify a root cause has been made and a corrective action (or correction, if appropriate) has been identified.

## **2.15 Failure To Provide Acceptable Car**

Should a participant fail to provide an acceptable corrective action report for a failed analyte, a warning notice with a deadline will be issued, reminding the participant of the consequences should an acceptable CAR not be provided. If the participant still fails to provide an acceptable CAR, the PT status for the affected analyte will be advanced to the next level (e.g., *Possible Suspension* to *Suspension*), adjustments will be made to the participant's scope of accreditation and the participant will be notified of the change.

## **2.16 Information Kept On Each PT Round**

The following information is maintained for each PT round:

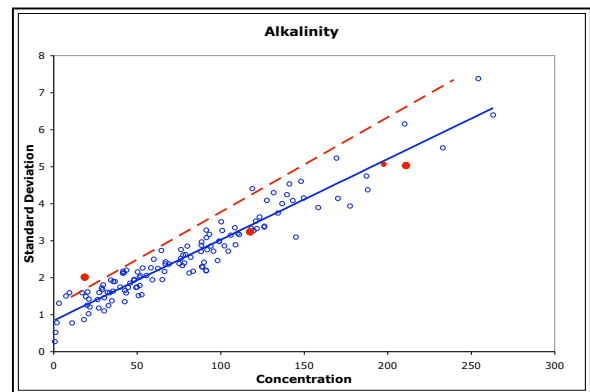
- Records of suitability of PT samples;
- Non-conformances, remedial actions, and corrective actions for ongoing PT samples;
- Instruction sheets for each round; and,
- Summary reports including Reference Value Report, Test Value Report, Proficiency Summary by Parameter, and the PT Round Comments Report.
- Confidential PT Report;
- Test Group Summary Reports;
- Reference value reports;
- Notices of Status Change; and,
- Blank CAR Reports.

### 3.0 CORRECTIVE/REMEDIAL ACTION FOR THE PT ROUND

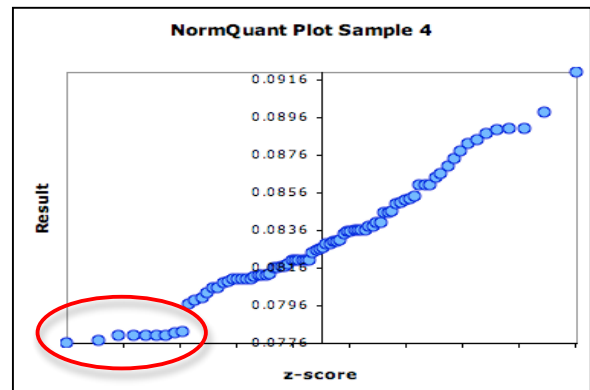
All aspects of the PT program fall under CALA's continuous improvement procedures (Q24-*Continual Improvement*). There are some procedures, because of the use of Collaborator and Reference laboratories, which are specific to the PT program.

Although section 2 details some default actions to take when some non-conformances are identified, further actions may be taken at the discretion of the PTM. This may entail, but is not limited to, the following:

- A plot of historic standard deviations against assigned value, including the active round, may be produced to see if there is an obvious explanation for the flag. For example, the plot to the right displays one sample that had an inter-laboratory standard deviation that was higher than the warning limit. However, as the concentration was at the low end of the scatter plot, the flag is likely the result the bias at either end of the range that is inherent in the establishing of the warning limits;



- A normal quantile plot may be produced. If there are some points that do not appear to fall along the trend line, they may be removed and the assigned values re-calculated.



If the PTM judges that proceeding as normal will have an impact on the integrity of the round the following steps are followed:

- If only one sample is affected, that sample is removed from the evaluation procedure for the affected PT round. If more than one sample is affected, actions may include, but are not limited to:
  - re-prepare the sample(s);
  - exclude the affected sample(s) from the evaluation procedure; or,
  - exempt the entire analyte from evaluation in the affected PT round.

- Proficiency *Testing Reports* will clearly identify when these deviations have been made. Whenever deviations are made, a notice is also included with the report explaining the rationale for modifying the evaluation procedure; and,
- The PTM will review and document the effectiveness of the corrective action during the next PT round.

### 3.2 Preventive Actions

Non-conformances or opportunities for improvements (see Q24 for details) can be identified at all steps during the PT round. For example, a large number of similar comments during web data entry may identify a problem with the samples that is not identified as part of the routine data examination.

Any one of these observations may trigger the generation of a corrective action form and subsequent root cause analysis and identification of corrective or preventive action by the Collaborator or by CALA.

As well, CALA receives feedback from several sources (e.g., Program Committee, workshops, presentations, training, etc.), all of which can trigger opportunities for improvement.

### 3.3 Correction Reports

Whenever a non-conformance is identified after the final Proficiency Testing Reports have been issued, the non-conformances are investigated as above and, if necessary, results are re-evaluated and revised Proficiency Testing Reports are issued. The reports are clearly indicated as being **Revised** reports and the cover page(s) of the revised report will include a brief statement about the reason for the revision.

Whenever corrections result in changes to a laboratory's PT or accreditation status, the necessary changes are also made to the Scope of Accreditation if applicable.

## APPENDIX I: PT EVALUATION PROCEDURE

The following procedure is used for each test group except C05B microbiology by presence/absence and The PCB aroclors in C06B, C08 and C35. The details for these can be found in Appendix II.

### AI.1 The Assigned Value ( $\bar{X}$ )

The default condition is to use the Robust Mean as described in Algorithm A of ISO 13528-*Statistical methods for use in proficiency testing by interlaboratory comparisons*.

### AI.2 The Inter-laboratory Standard Deviation (stdev)

The default condition is to use the Robust standard deviation as described in Algorithm A of ISO 13528-*Statistical methods for use in proficiency testing by interlaboratory comparisons*. Care must be taken when using this procedure for small data-sets containing numerous identical values. This will result in an unreasonably low estimation of the standard deviation. When this is observed, the arithmetic standard deviation will be used.

### AI.3 The Regression Equation Standard Deviation (s!)

The regression equation standard deviation is estimated using the regression equations estimated from historic studies (see PT15-05-CALA PT Program - *Regression Equations*) or established as fixed limits based on regulatory data quality objectives and analytical capabilities.

$$s! = m \times \bar{X} + b$$

where  $m$  = slope of regression equation

$\bar{X}$  = assigned value

$b$  = intercept of regression equation

### AI.4 The Assigned Deviation (s)

The assigned deviation used to calculate the z-score is determined as follows:

- The regression equation standard deviation  $s!$  is used if this number is higher than the inter-laboratory standard deviation (*stdev*);
- The inter-laboratory standard deviation (*stdev*) is used if this value is higher than the regression equation standard deviation.

### AI.5 z-Scores

The outliers and non-detects are added back into the database in order to prepare the *Final PT Reports* for all the participants.

The assigned value is rounded to the number of significant figures and decimal points that will appear on the final report. The assigned deviation is rounded to the number of significant figures and decimal places that will appear on the final report.

The z score and absolute z score for each reported result is calculated as:

$$z \text{ score} = \frac{(x_i - \bar{X})}{s}$$

$$\text{absolute } z \text{ score} = \frac{|(x_i - \bar{X})|}{s}$$

If a laboratory has reported its RDL during web-data-entry the following equations are used:

$$z \text{ score} = \frac{(x_i - \bar{X})}{\sqrt{s^2 + (RDL/3)^2}}$$

$$\text{absolute } z \text{ score} = \frac{|(x_i - \bar{X})|}{\sqrt{s^2 + (RDL/3)^2}}$$

Where  $x_i$  = reported result  
 $\bar{X}$  = assigned value  
 $s$  = assigned deviation  
 RDL = the reporting detection level

**Note:** The RDL option is not available for test groups, C04B, C05, C11, C12, C13, C15 or C20.

**Exception 1-Very High z Scores:** If the calculated z score is > 6.6 or < -6.6 then it is set to 6.6 or -6.6 respectively.

**Exception 2-Non-detect Values:** If the reported value is non-detect, the following rules apply:

- If reported non-detect value is less than the assigned value, use the non-detect value to calculate a z score;
- If reported non-detect value is greater than the assigned value, then:
  - set z score at 2 if the PT analyte/test group is only offered in one concentration range;
  - set z score at 2 if the PT analyte/test group is the high range of a two range PT test group (e.g., C02B); and,

- set z score at 3 if the PT analyte/test group is the full range or low range of a two range PT test group (e.g., CO2A).

**Note:** If a laboratory reports a result lower than the reported RDL, this will be treated as a non-detect and evaluated accordingly.

**Exception 3 - Greater than value:** For microbiology, an accurately reported greater than value will be assigned a z score of 2. For all other greater than values, the greater than sign will be ignored and the value evaluated as a normal result.

**Exception 4 - No Result Reported:** If no result is reported, or if a chemistry result is reported as zero, the z score is set to 6.6.

## AI.6 Calculate Composite PT Score

The average absolute z score is calculated for each test group/analyte combination as

$$avgz = \frac{\sum absolutezscore}{N}$$

where absolute z score = score as calculated above

N = number of samples per test group (generally 4).

The composite score for each test group/analyte is calculated as

$$PT\ Score = 100 + (-15 * avgz)$$

where *avgz* = average absolute z score as calculated above.

## AI.7 Estimating and Flagging Bias

Biases are identified using the rescaled z score procedure. Calculate the rescaled z score as

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

where z = the z score as calculated above

N = the number of samples in the test group (generally 4)

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

- RSZ > -2 and < 2      no flag assigned;
- RSZ > 2                      H (High);

- RSZ > 3                      VH (Very High);
- RSZ < -2                    L (LOW); and,
- RSZ < -3                    VL (Very Low)

## AI.8 Standard Uncertainty of the Assigned Value:

The standard uncertainty of the assigned value is estimated as:

$$u_x = 1.25 \times stdev / \sqrt{N}$$

Where stdev = inter-laboratory standard deviation as estimated above.

N = number of participants (outliers removed).

## AI.9 Interpretation of PT Results

The PT round specific Acceptable/Unacceptable status for each test group/analyte combination is assigned as:

- PT score  $\geq$  70 = Acceptable;
- PT score < 70 = Unacceptable.

## AI.10 Custom Reports

If any of the evaluations are modified (e.g., eliminate one sample, eliminate entire analyte, etc.) the custom report clearly identifies the modified evaluation. The reason for the modification is also included in a PT Notice provided with the final report.

Custom PT reports are emailed to participants as pdf files. These reports include the Proficiency Testing Report, a cover letter, a notice of proficiency testing changes, if applicable, and Corrective Action Forms for any unacceptable PT scores.

In addition to the mailed reports, electronic csv files are emailed to participants. These csv files contain the same information that is included in the custom report.

Each custom report contains information about the participant as well as,

Sample ID	PT code
Appendix number	Analyte
Method	Units
Assigned Value	Reported Value
Standard uncertainty of assigned value	s Value
z Score	Bias flags
Composite PT score	
Acceptable/Unacceptable status	



The *Test Group Summary Reports* (PT35) are uploaded to the CALA web site for participant access. As well, any PT Notice needed and PT Challenge Sample reports, when applicable, are also uploaded.

## **APPENDIX II: PT EVALUATION FOR MICROBIOLOGY (PRESENCE/ABSENCE) AND PCB AROCLORS**

### **AII.1 Microbiology by presence/absence**

For the C05B microbiology samples, an unacceptable evaluation is assigned for any false positive or false positive.

### **AII.2 PCB Aroclors**

PCB aroclors are evaluated as a combination z-score and presence/absence procedure. Each of the four samples in the test group is spiked with a single aroclor. For the aroclors that are not spiked into the sample, a threshold concentration is estimated as a fraction of the spiked aroclor concentration. For each aroclor, if a laboratory reports a false positive at a concentration above the relevant threshold value, then the aroclor is assigned an *UNACCEPTABLE* evaluation regardless of any calculated z-scores. If there are no unacceptable false positives, then the PT score is evaluated based on the z-score(s) of the sample(s) that were spiked with the aroclor.