

# Interpretation of Thistle EQA Reports on the New Portal System and Statistical Protocol

---

**Purpose** - The purpose of this document is to help labs gain the maximum information from Thistle QA reports and understand the Statistical Protocol applied to the schemes on our New Portal System. Please note that we are still in transition with moving all our schemes to Portal so the below information does not apply to our Differential Slide and Microbiology Scheme.

## Contents

Important information about the Statistical Protocol changes.....	3
Interpretation of the Analyte Report.....	3
Cover page .....	3
Scheme Details.....	4
Assay Details .....	4
Cumulative Summary .....	4
Performance Assessment by Analyte Type: Quantitative .....	5
Your Result Assessment.....	5
Statistical Data .....	7
Assessment Statistics .....	8
Performance Score histogram .....	10
Trend Charts.....	10
Method Summary .....	11
Instrument Summary .....	11
Performance Assessment by Analyte Type: Qualitative .....	12
Result Assessment .....	12
Assessment grading guide .....	12
Qualitative Assessment Statistics .....	13
Last 12 Qualitative Assessments:.....	13
Performance Assessment by Specific Analyte Type: Troponin T Categories.....	14
Result Assessment .....	14
Qualitative Assessment Statistics .....	14
Assessment Graph .....	15

Last 12 Qualitative Assessments:.....	15
Performance Assessment by Specific Analyte Type: HIV Qualitative .....	15
Result Assessment .....	15
Assessment Statistics .....	16
Last 12 Qualitative Assessment .....	17
Performance Assessment by Specific Analyte Type: Titres .....	17
Statistical Protocol .....	18
Data Analysis and Performance Assessment .....	18
Understanding the Statistical Data on your Report.....	19
Procedure for calculating robust statistics .....	20
Robust mean (median) .....	20
Robust Standard Deviation .....	21
Removal of errors and blunders .....	22
Estimated Standard Uncertainty of the assigned value.....	22
Sample ranges around the CV Cut off.....	23
Acceptable but outside 2 SDs .....	23
Inside 2 SDs and A for Acceptable BUT consistently biased .....	24
Important General Requirements for ISO 17043 covered in our reports.....	24
Support: .....	25

## Important information about the Statistical Protocol changes

Chauvenet's Criterion will no longer be used to perform statistical evaluation; instead robust statistics will be used as described in this document. As a result of this change in the statistical protocol we have not carried over any of your historic data.

On your request, we would be happy to re-send any reports that you might need from our legacy system. Please note these would be in our legacy report format and not in the new format of the reports. As per our data retention policy - electronic raw data is kept permanently but may only be retrievable in pdf format for up to 5 years back depending on technology and system changes. Any data that is not retrievable in an PDF report format will still be retrievable but in an excel format.

You will only receive an Analyte and a group report, our legacy cumulative reports will now be included in the Analyte report and the group report will include any laboratories that did not submit any results. Groups have been created according to the legacy system information, contact us to customise your group report members as required.

## Interpretation of the Analyte Report

### Cover page

Report Cover Page will contain the following information:

- Thistle Logo's
- Your Lab ID and Lab Name
- The title of the Report
- Round with Round Description in Brackets
- Issue Number
- Issue Date
- Thistle's Contact details
- Accreditation Body Logo
- Page numbers

### How do I identify my Laboratory on the Report?

- Each laboratories Lab ID and name will be displayed on the cover page of your report. All of the pages thereafter will have the Lab and laboratory name in the header at the top of the page.

### How do I know what scheme and round the report is for?

- The round followed by the scheme name is displayed on the cover page. All of the pages thereafter will have the round and the scheme displayed in the header on top of the document.

## Scheme Details

### Aims of Scheme

This section outlines the Technical aims of the scheme

### Sample details

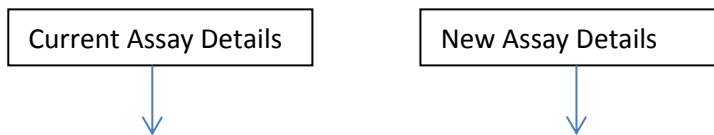
This section outlines the sample specifics

### Quality Control

This section outlines the ISO17043 requirements

## Assay Details

Your assay details are displayed in a table with the method profile column being your current assay details. Should there be changes to your assay details/method profile – there will be an additional column labelled “New method Profile Effective from this Round”. You can use this as a quick reference to see what you are currently registered for and what method profile you are registered to.



### Assay Details

Details	Method Profile	New Method Profile Effective from this round
Analyte:	BNP	BNP
Primary Method:	Siemens Atellica IM	
Instrument:	Siemens Atellica Solution	
Supplier/Kit Code:	Siemens/Bayer/Technicon	
Slope/Intercept:	1.000/0.000	
Unit:	ng/L	pmol/l
Reference:	Siemens Atellica Solution Siemens Atellica IM	

## Cumulative Summary

A cumulative summary of your last six performance scores, results submitted and assessed results. Clinical acceptance is indicated as A = Acceptable and P = Poor, your performance score is indicated in brackets.

Cumulative Summary

Round\_ (Deadline date)

Statistics

Analyte	THLC011_6 (01/99)	THLC011_7 (01/99)	THLC011_8 (01/99)	THLC011_9 (01/99)	THLC011_10 (01/99)	THLC011_11 (01/99)
High Sensitivity CRP	A(-0.61)	A(0.18)	A(0)	A(0)	A(-0.1)	A(0)

Total Number of Results Submitted: 6 Out of: 6 Results = 100.0%  
Total Number of Results with Satisfactory performance: 6 Out of: 6 = 100.0%  
Total Number of Numeric Results Acceptable: 6 Out of: 6 = 100.0%  
Total Number of Numeric Results Not Assessed: 0 Out of: 6 = 0.0%

## Performance Assessment by Analyte Type: Quantitative

### Your Result Assessment

Analyte: BNP

Instrument / Method: Siemens Atellica Solution | Siemens Atellica IM

Results

Raw Result	Conversion Factor	Result	Unit	Stats By	Perf Score	Score Type	% Diff	Clinical Acceptance
85.89	3.4602	24.822	pmol/l	Method	0.638	zPrime	4.294	ACCEPTABLE
85.89	3.4602	24.822	pmol/l	Instrument	0.638	zPrime	4.294	ACCEPTABLE

#### Analyte

The analyte to which the statistical results apply

#### Instrument/Method

The instrument and method profile that your lab is registered against.

#### Raw Result

This is the result that was captured on the Result submission screen or submitted to Thistle to capture.

#### Conversion Factor

If your result is in a different unit to the Thistle default unit the conversion factor we are using to convert your result is displayed in this column. If you are using the Thistle default unit the conversion factor column will be populated with a 1.

#### Result

This is the result we have used to assess your performance as the conversion factor would have been applied to your raw result.

#### Unit

The Thistle default Unit

#### Stats By

This will indicate either Method or Instrument on each line. If it indicates Method the stats is based on your method and if it indicates instrument then the stats is based on your Instrument.

#### Perf Score

Your performance score.

Performance scores are colour coded to indicate their assessment. Green indicates 'Satisfactory' for absolute scores  $\geq 0$  and  $\leq 2$ . Orange indicates 'Questionable' for absolute scores  $> 2$  and  $< 3$ . Red indicates 'Unsatisfactory' for absolute scores  $> 3$ .

Performance scores will not be given for the following:

- For qualitative results, where satisfactory performance is based on the participants reporting the same result as the assigned result. E.g. detected, not detected. For these results, colour coding of green (satisfactory) or red (unsatisfactory) will apply.
- For results of zero; such a result is not normally appropriate and should not be reported, the result should be reported as less than the detection limit rather than zero
- For quantitative results where the analyte under test is present in the test material but participants report zero results or greater than results. In these cases, it is not possible to allocate a performance score and participants should assess their performance based on the assigned value and satisfactory range given.
- For quantitative results where the analyte under test is present in the test material but participants report a 'less than' value. In these cases, it is not possible to allocate a numeric performance score, however, where the 'less than' value reported is  $< (AV-3*SDPA)$  the 'less than' value will be assessed as unsatisfactory (red colour coding), where the less than value reported is between  $< (AV-3*SDPA)$  and  $< (AV-2*SDPA)$ , or  $> (AV+2*SDPA)$  the assessment will be questionable (orange colour coding) and it is recommend that you assess whether the method used is fit for purpose, and where the less than value reported is between  $(AV-2*SDPA)$  and  $(AV+2*SDPA)$  a satisfactory assessment (green colour coding) will be given as such results are deemed to be consistent with the assigned value.

In some cases, performance scores may not be provided or may be provided but with colour coding suspended (indicating that scores need to be interpreted with caution). For example:

- For small data sets where less than 5 results have been submitted and the assigned value is derived using a consensus value from the participants' results. In these circumstances, there may be increased uncertainty of the assigned value, given the low number of participants, and performance scores will be given for information only.
- In cases where the distribution of the results gives cause for concern e.g. bi-modal data sets. These circumstances will be dependent on the statistical design that is in place.
- If the assigned value falls below a concentration threshold (only applies to some schemes).
- In these or similar circumstances, further explanation as to the reasons for suspension of performance scoring or colour coding, and on the interpretation of results, will be given in the report

## Score Type

Two types of performance scores are currently possible z score and z' score (z prime). This field indicates the type of score in use. By default the score type is Z score.

The participant's result,  $x$ , is converted into a performance score using the following formula:

Z score formula =  $\text{Result-Assigned value} / \text{SDPA}$

Where:

SDPA = Standard Deviation for Proficiency Assessment

For small data sets, there will be increased uncertainty around the assigned value if derived from the consensus value of participants' results. In such cases, z performance scores may not be provided, or may be given for information only.

The Z score expresses performance in relation to the assigned value and the standard deviation for proficiency assessment (SDPA).

A z performance score of 2 represents a result that is a distance of 2 x SDPA from the assigned value.

A z' performance score (z-prime) incorporates the standard uncertainty of the assigned value and is calculated as follows:

$$Z' \text{ Score Formula} = \frac{\text{Result} - \text{Assigned value}}{\text{SQRT}(\text{SDPA}^2 + \text{UxAV}^2)}$$

Where:

SDPA = Standard Deviation for Proficiency Assessment

Ux = standard uncertainty of the assigned value X

A z' performance score is interpreted in exactly the same way as a z performance score, ≤2 is satisfactory, >2 but <3 is questionable and ≥3 is unsatisfactory.

### % Diff

The difference between your result and the Assigned Value is calculated. This figure is then divided by the Assigned Value and multiplied by 100, giving the percentage difference that your result is from the Assigned Value.

### Clinical Acceptance

Sample Performance shows your performance for the current sample, and will be expressed by either an A for Acceptable or P for Poor. The expression of the performance as Acceptable or Poor does not apply for the HIV, Pregnancy and 5 Part differential programmes.

### Statistical Data

Assessment Statistics	Method	Instrument	Result Statistics	Method	Instrument
Assigned Value	23.800	23.800	Number of Results	7	7
Assigned Value Derived By	each primary method	Instrument/Primary Method	Number of Excluded Results	0	0
Uncertainty of Assigned Value	0.685	0.685	Mean	24.704	24.704
SDPA	1.449	1.449	Median	23.800	23.800
Acceptable CV	20.0	20.0	Standard Deviation	2.1009	2.1009
Satisfactory Range	20.594 to 27.006	20.594 to 27.006	Robust SD	1.4493	1.4493
Satisfactory z/z score	85.7%	85.7%			
Questionable z/z score	0.0%	0.0%			
Unsatisfactory z/z score	14.3%	14.3%			

## Assessment Statistics

All statistics are calculated in accordance with ISO 13528 and use robust statistics which mitigate the effect of outliers on the data set. For more in depth information review the General Protocol.

### Assigned Value

The assigned value is derived using the Robust Mean (Median) which calculated following the removal of gross errors and blunders from the original dataset.

For quantitative data, gross errors or blunders are removed from the data by removal of any results that are greater than the assigned value  $\pm 5 \times$  SDPA. These results are not used in the final calculation of the assigned value and other summary statistics and will be included in the number of 'Excluded Results'. All results, including excluded results, will be given a performance score.

### Assigned Value Derived By

An indication on how the statistical data is grouped together to form your assessment group by method profile derivatives.

This is a replacement for the Stat levels that Thistle used in TQA.

You will see one of the below combinations on each analyte:

Method	Instrument
All Primary Methods Comparable	Split by Instrument
Split by primary method/mode combination	Split by Instrument/Primary Method Mode
Split by primary method	Split by Instrument/Method Combination

### Uncertainty of Assigned Value

The assigned value has a standard uncertainty ( $u_x$ ) that depends upon the method used to derive the assigned value. When the assigned value is determined by the consensus of participants' results, the estimated standard uncertainty of the assigned value can be calculated by:

$$u_x = 1.25 \times \text{Robust standard deviation}/\sqrt{n}$$

Where  $n$  = number of results

### SDPA

The SDPA expresses the acceptable difference between the laboratory result and the assigned value. An acceptable  $z$  performance score represents a result that does not deviate from the assigned value by more than twice the SDPA.



## Acceptable CV

If your result as a %D is within this percentage from the mean result, you will be regarded as “acceptable’ for that result i.e. if your result is less than or equal to the Acceptable CV. These figures come from a variety of sources, such as CLIA’88 (the USA regulation describing satisfactory performance on EQAs) with elements of Biological Variation (BV). From these pieces of information we have taken local advice and created our own South African set of acceptable performance standards.

At a specific concentration certain analytes have a variation of the standard acceptable CV for that Analyte.

This will be indicated as Acceptable CV (+-Value AV Cut off Assigned Value)

An example of this would be 36 % (+-7 AV 20) for Total Bilirubin. This will mean that if the assigned value for this sample is less than 20 you will still be acceptable if your result is in the range of 13 -27. This is calculated by subtracting 7 from 20 for your minimum acceptable value and adding 7 to 20 for the maximum acceptable value

## Satisfactory Range

This displays the minimum and maximum values that are within 2\*SDPA from the assigned value.

## Satisfactory z/z score

The percentage of satisfactory performance score.

## Questionable z/z score

The percentage of questionable performance score.

## Unsatisfactory z/z score

The percentage of unsatisfactory performance score.

## Result Statistics

### Number of Results

This is the number of results that is part of your assessment group, expressed for both Instrument and Method.

### Number of Excluded Results

This is the number of results excluded by the statistical process. I.e. Gross errors/blunders and < or > values.

### Mean

Average Value of the data set

### Median

Middle Value of the data set, when sorted ascendingly, following the removal of excluded results.

## Standard deviation

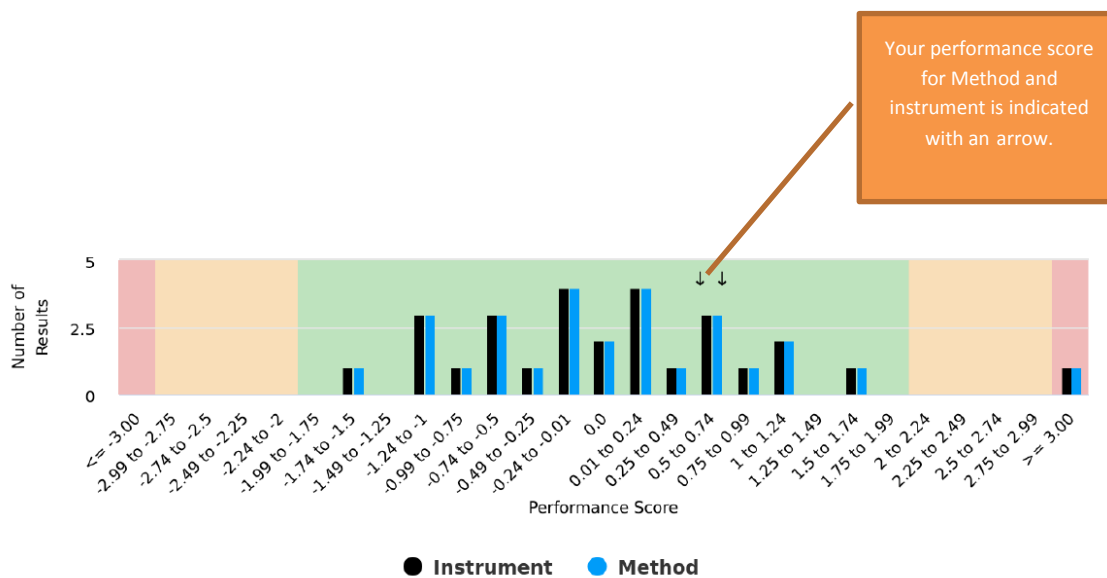
A measure of the amount of variation in the reported results

## Robust SD

The normalised Median of Absolute Deviations ( $MAD_E$ ) from the sample median is used as a robust standard deviation.

## Performance Score histogram

The performance score histogram shows the dispersion of both the method and instrument performance scores of all participants in your method/instrument group.



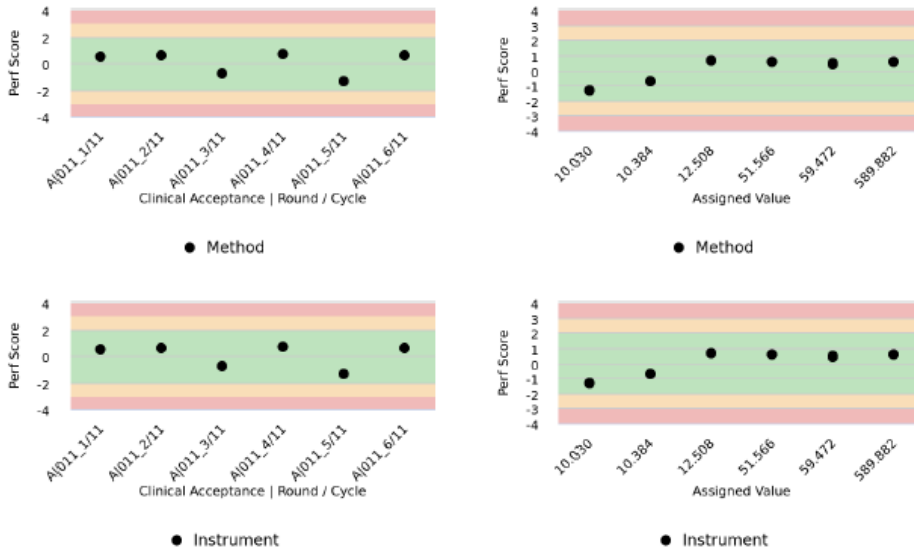
## Trend Charts

### Performance Score Trend Chart by Clinical Acceptance/Round

Your performance score for the last 12 rounds is plotted against the clinical acceptance and round for both the instrument and method data set.

### Performance Score Trend Chart by Assigned Value

Your performance score for the last 12 rounds is plotted against the assigned value for both the instrument and method data set.

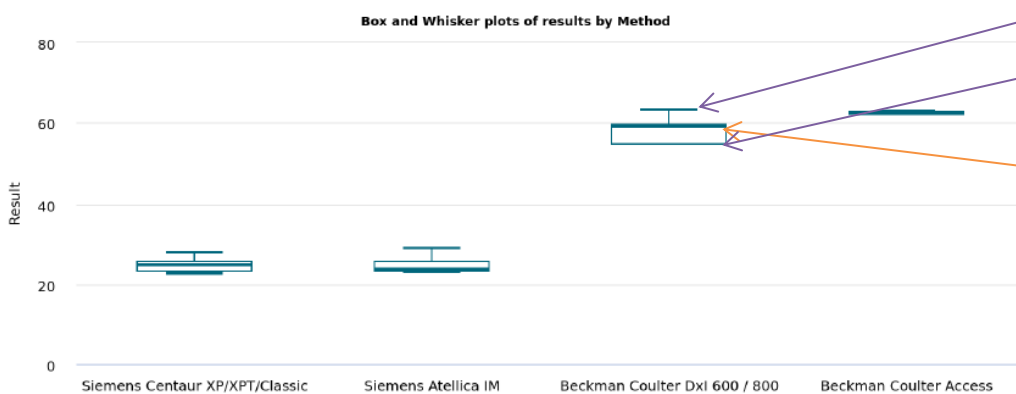


## Method Summary

This table shows you a data summary of the methods reported for this analyte followed by a Box and Whisker graph.

This type of graph is used to show the shape of the distribution, its central value, and its variability. In a box and whisker plot: the ends of the box are the upper and lower quartiles, so the box spans the interquartile range, the median is marked by a vertical bold line inside the box.

Method Summary					
Method	Number of Results	Median	RSD	Range	Sat %
Siemens Centaur XP/XPT/Classic	15	24.94	1.626	22.362 to 27.631	100.0
Siemens Atellica IM	7	23.8	1.449	22.823 to 28.986	85.7
Beckman Coulter Dxl 600 / 800	4	58.87	3.656	54.52 to 63.22	100.0
Beckman Coulter Access	2	62.35	0.43	62.06 to 62.64	100.0



## Instrument Summary

This table shows you a data summary of the instruments reported followed by a Box and Whisker graph.

This type of graph is used to show the shape of the distribution, its central value, and its variability. In a box and whisker plot: the ends of the box are the upper and lower quartiles, so the box spans the interquartile range. The median is marked by a vertical bold line inside the box.

Author: R.Otto /T.Ramjee

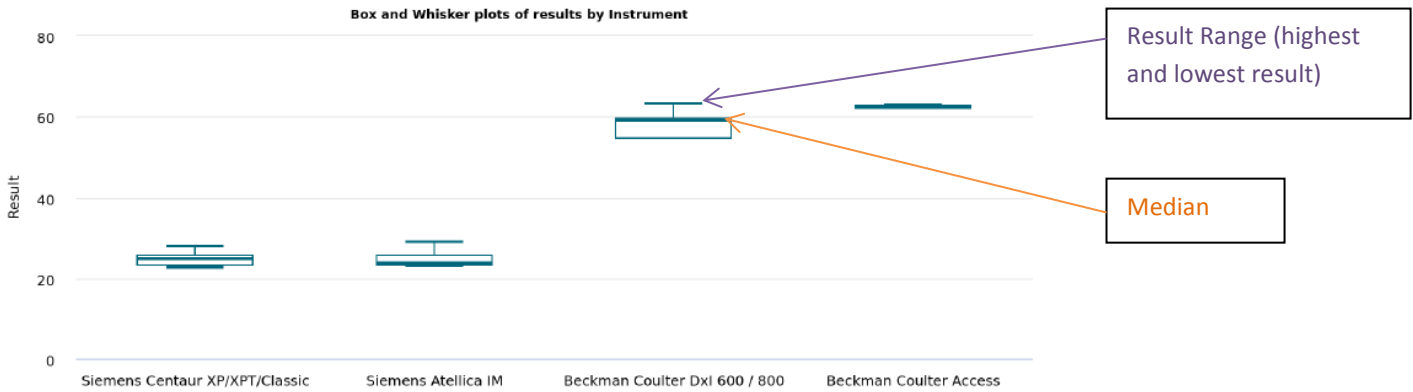
Authorised by: R.Otto

T:\Quality Management System\Quality Documents\Customer Information Documents-CID

Issue date: 28/01/2020

Instrument Summary

Instrument	Number of Results	Median	RSD	Range	Sat %
Siemens ADVIA Centaur XP/XPT/Classic / Siemens Centaur XP/XPT/Classic	15	24.94	1.626	22.362 to 27.631	100.0
Siemens Atellica Solution / Siemens Atellica IM	7	23.8	1.449	22.823 to 28.986	85.7
Beckman, Dxl 600 /Dxl 800 / Beckman Coulter Dxl 600 / 800	4	58.87	3.656	54.52 to 63.22	100.0
Beckman Access / Beckman Coulter Access	2	62.35	0.43	62.06 to 62.64	100.0



## Performance Assessment by Analyte Type: Qualitative

### Result Assessment

Your result will be listed here with assessment grading colour

Result
Positive

### Assessment grading guide

Interpretation	Colour coding
Satisfactory result	Green
Questionable result	Amber
Unsatisfactory result	Red

## Qualitative Assessment Statistics

### Qualitative Assessment Statistics

Assessment Statistics		Result
Split by each primary method	Assigned Value	Positive
	Assigned Value Derived By	each primary method
	Number of Results	1
	Satisfactory performance %	100.0%

#### Assigned Value

The assigned value is derived by consensus.

We require 80% of lab results to agree before we have a consensus result. If the report says “No Consensus” it means that less than 80% of labs did not agree. In this case the lab can act according to their procedures and/or contact Thistle for any assistance needed. If the report says “Too few” it means that there were less than 5 results and thus it has not been possible to achieve consensus.

#### Assigned Value Derived By

An indication on how the statistical data is grouped together to form your assessment group by method profile derivatives.

#### Number of Results

This is the number of results that is part of your assessment group, expressed for both Instrument and Method.

#### Satisfactory performance %

Percentage of results that is satisfactory

\*Should there be unsatisfactory performance, the percentage of unsatisfactory results will display.

## Last 12 Qualitative Assessments:

A below assessment summary shows the last 12 assessments, your assessment is indicated with a prefix and your assessment grading colour.

**Last 12 Qualitative Assessments**

Method	THLC012 1
Biomerieux Vidas/miniVidas/Vidas 3   bioMerieux Vidas	S

S = Satisfactory | Q = Questionable | U/S = Unsatisfactory | < = Too few results | NC = No Consensus | NA = Not Assessed | X = No result submitted

## Performance Assessment by Specific Analyte Type: Troponin T Categories

### Result Assessment

Your result will display in the category range you have selected with your assessment grading colour.

Result
Medium Risk (0.03 - 0.1 ng/ml or 50-100 ng/l)

## Qualitative Assessment Statistics

**Qualitative Assessment Statistics**

Assessment Statistics		Result
All data comparable	Assigned Value	Medium Risk (0.03 - 0.1 ng/ml or 50-100 ng/l)
	Assigned Value Derived By	All data
	Number of Results	18
	Satisfactory performance %	83.3%
	Unsatisfactory performance %	16.7%

### Assigned Value

The assigned value is derived by consensus.

We require 80% of lab results to agree before we have a consensus result. If the report says “No Consensus” it means that less than 80% of labs did not agree. In this case the lab can act according to their procedures and/or contact Thistle for any assistance needed. If the report says “Too few” it means that there were less than 5 results and thus it has not been possible to achieve consensus.

### Assigned Value Derived By

For all Category analytes it will be all data comparable

### Number of Results

This is the number of results that is part of your assessment group, expressed for both Instrument and Method.

Satisfactory performance %

Percentage of results that is satisfactory.

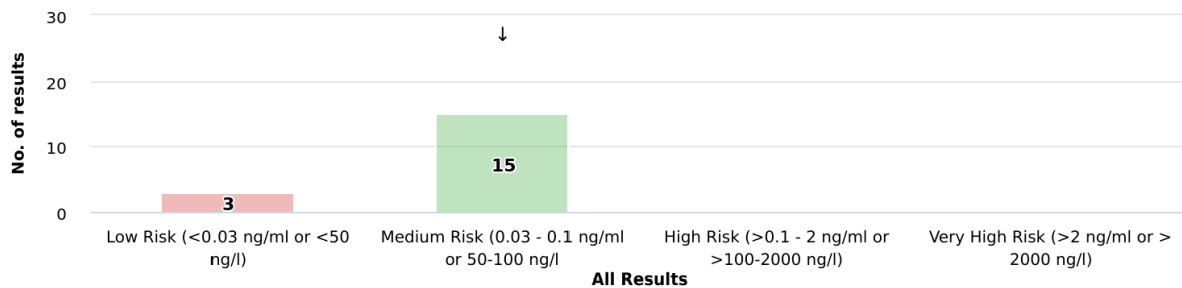
\*Should there be unsatisfactory performance, the percentage of unsatisfactory results will display

## Assessment Graph

The Assessment graph shows all the categories of result submission and how many results were submitted in each category.

Each category displays the applicable assessment grading.

The arrow indicates your result category.



## Last 12 Qualitative Assessments:

A below assessment summary shows the last 12 assessments, your assessment is indicated with a prefix and your assessment grading colour.

### Last 12 Qualitative Assessments

Method	THLC012 1
Roche Cobas h 232   Roche Cardiac Reader	S

S = Satisfactory | Q = Questionable | U/S = Unsatisfactory | < = Too few results | NC = No Consensus| NA = Not Assessed | X = No result submitted

## Performance Assessment by Specific Analyte Type: HIV Qualitative

### Result Assessment

Will be displayed as following with the colour grading fill based on your performance:

Green – Satisfactory Performance:

Result
Positive

\*We require 80% of the submitted results to agree per instrument before we achieve a Consensus result.

Red – Unsatisfactory Performance

Non fill – Performance not assessed as the grading criteria has not been met:

Result
Positive

\*Performance is not graded in the event that <5 results are submitted or consensus is not achieved.

## Assessment Statistics

### Qualitative Assessment Statistics

Assessment Statistics		Result
Split by each primary method	Expected Result	Positive
	Assigned Value	Positive
	Assigned Value Derived By	each primary method
	Number of Results	19
	Satisfactory performance %	100.0%

**Split by:** Indicates the Split by each Primary Method(this is actually the instrument)

**Expected Results** – this is a reference value listed as to what the consensus is expected to be.

**Assigned Value:**

Assigned value is derived from 80% of the consensus for the results submitted per instrument, meaning if 80% reported Positive, then the consensus is positive. In the event that 80% consensus is not archived, the assigned value will be listed as too few and the performance is not assessed.

The minimum result count for grading is 5, Too Few will be listed in the event that <5 results are submitted.

**Assigned Value Derived By:**

HIV Qualitative will be each primary method

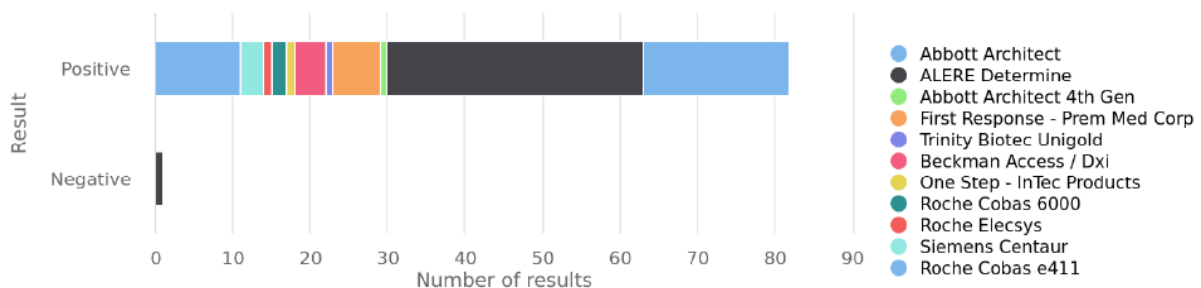


**Number of results** – number of results submitted

**Satisfactory Performance %** - this is the percentage of results considered satisfactory for the instrument

**Unsatisfactory Performance %** - this is the percentage of results considered unsatisfactory for the instrument

Graphical Representation of the total number of results returned across all instruments:



## Last 12 Qualitative Assessment

This will outline your performance across 12 samples. Performance is colour graded with the same colour grading as your result.

Last 12 Qualitative Assessments

Method	THHS054_1
Abbott Architect	S

S = Satisfactory | Q = Questionable | U/S = Unsatisfactory | < = Too few results | NC = No Consensus| NA = Not Assessed | X = No result submitted

## Performance Assessment by Specific Analyte Type: Titres

Titre graphical representation will contain the following logic:

The chart should fix to 5 bins and allocate data to the bins as follows:

1. The central bin should be the most common results value report.
2. There should then be 2 bins either side of the most common value based on their position in the result field value order.
3. When 'Not detected' is the most common value there should be 2 empty bins to the left hand side because 'Not detected' occupies result field value order position 1.
4. The first and fifth bin should be used to collate all results that do not fit in the second or fourth bin. The bin (series name) for these should be the name of the second or fourth bin prefix with a < or > symbol respectively.

## Statistical Protocol

### Data Analysis and Performance Assessment

Thistle QA has a wide range of schemes, which may include qualitative, quantitative, semi-quantitative and interpretive tests. Different approaches to data analysis may therefore be used, the most common approaches being described below. Further information on the statistical approach for specific schemes is also provided in the Scheme Descriptions and Scheme Reports.

The advantages of using a performance score are:

- Results can be expressed in a form that is relatively easy to interpret and understand
- Results can be summarised in graphical or tabular form to depict overall performance
- A performance score allows participants to directly compare their own result with others
- If consistent statistical values are applied, a performance score enables participants to monitor trends in their own performance, over time.

When reviewing results, participants should take into account the methods used to analyse the data and to assess performance, and should review their performance in context, taking into account performance of the whole dataset.

#### Qualitative schemes

For qualitative tests, participant results will be compared against the intended result, also called the assigned value, based on formulation or expert assessment. A result which is the same as the assigned value is considered satisfactory. This approach is also used for quantitative tests when the target analyte is absent and for semi-quantitative tests where the assigned value may be a range of results. For interpretive schemes where the result is subjective rather than quantifiable, a model answer produced by appropriate experts will be published in the report.

#### Quantitative schemes

For quantitative data, participants are assessed on the difference between their result and the assigned value (see 5.4); with this difference being represented by a performance score called a z or z' (z prime) score

#### Setting assigned values

The assigned value is the value selected as being the best estimate of the 'true value' for the parameter under test. The method used to determine the assigned value may vary depending upon the particular scheme and test parameter, and is detailed in the relevant scheme description, along with details of the traceability in each case.

For quantitative tests, all assigned values are derived in accordance with ISO 13528. Where it is appropriate, practicable and technically feasible the assigned value will be derived through formulation (or occasionally through the use of a certified reference material) to provide metrological traceability; the associated uncertainty

of the value can therefore be estimated. However, in most cases it will not be possible to use formulation or certified reference materials to set the assigned value and a consensus value will be the only practicable and technically feasible approach to use. When the assigned value is determined from the consensus value of participant results, or from expert laboratories, robust statistical methods are used for calculation of the consensus value, details of which are outlined in this documents. The uncertainty of the assigned value is then estimated as described.

### Precision of Reported Statistics

You may notice that if you generate your own statistics many tools (e.g. Excel) will return Average, for example, too many more decimal places than you will see on your report. Thistle QA policy, as far as possible, is to display on your reports the average (or mean) to one decimal place higher than the typical precision of results submitted to Thistle QA for that particular analyte. We display the standard deviation to two decimal places higher than the typical precision of submitted results for that analyte. As an example, plasma Sodium is typically reported in whole integer values, e.g. 137, 147 mmol/l, on our reports you will see the average displayed as 134.8 as an example and 0.43 as the Standard Deviation.

## Understanding the Statistical Data on your Report

Once a PT round has closed, the results will be analysed and the assigned value determined for each analyte, according to the criteria provided in the Scheme Description. Information regarding the traceability of each calculated assigned value is also provided in the Scheme Description.

**Z Score** - For quantitative data, the participant's result, x, (or log<sub>10</sub> x for microbiological data) is converted into a z score using the following formula;

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

X = Assigned value

SDPA = Standard deviation for proficiency assessment

**Z' Score** -For quantitative data, the uncertainty of the assigned value is calculated to ensure that it would have a negligible effect on participants' performance scores. If the uncertainty of the assigned value is greater than 0.3 x SDPA, then this is not considered negligible. In this situation, a z' (z prime) performance score is automatically calculated rather than a z score, in order to take account of the measurement uncertainty of the assigned value. The z' score is calculated using the following formula;

$$z' = \frac{(x - X)}{\sqrt{SDPA^2 + u(x_{pt})^2}}$$

X = Assigned value

SDPA = Standard deviation for proficiency assessment

$u(x_{pt})$ = Uncertainty of the assigned value

$$\text{Expanded SDPA} = \sqrt{\text{SDPA}^2 + u(x_{pt})^2}$$

Trend graphs will use a mixture of z and z' scores, i.e. the 'performance score' for the round.

### Interpreting results

For qualitative or semi-quantitative results, laboratories reporting the assigned result or range of results will be considered correct, and therefore have satisfactory performance.

For the purposes of performance assessment for a single round, z and z' scores are interpreted as follows:

z/z' score	Interpretation	Colour coding
$ z  \leq 2.00$	Satisfactory result	Green
$2.00 <  z $ and $< 3.00$	Questionable result	Amber
$ z  \geq 3.00$	Unsatisfactory result	Red
No score given	See below	No colour coding

For small data sets (generally with less than 8 results) there will be increased uncertainty around the assigned value if using consensus values from participants' results. For those analytes that use a formulation or reference value as the assigned value and a fixed fit for purpose SDPA z scores will be provided. Where the assigned value and/or SDPA is based on participant results, performance scores will be given for information only. For data sets with very limited results or where the spread of results is large, z scores may not be provided.

For quantitative data, gross errors or blunders are removed from the data by removal of any results that are greater than the assigned value  $\pm 5 \times$  SDPA. These results are not used in the final calculation of the assigned value and other summary statistics and will be included in the number of 'Excluded Results'. All results, including excluded results, will be given a performance score.

## Procedure for calculating robust statistics

### Robust mean (median)

The consensus value can be calculated using the robust mean of all participant results. In LGC Standards PT schemes the robust mean used is the median. If the data, where there are an odd number of results are arranged in order of magnitude ( $x_1, x_2, \dots, x_n$ ) the median is the central member of the series, i.e. there are equal

numbers of observations smaller and greater than the median. Where there is an even number of results, the median is the average of the middle pair of numbers within the series. For a normal distribution the mean and median have the same value. The median is more robust, in that it is virtually unaffected by extreme values.

## Robust Standard Deviation

In LGC Standards PT schemes the normalised Median of Absolute Deviations (MADE) from the sample median is used as a robust standard deviation.

$MAD = \text{median} ( |x_i - X| \text{ } i = 1,2,\dots,n)$  where  $n$  = number of results

For example:

Data (g)	5.6	5.4	5.5	5.4	5.6	5.3	5.2
Ordered Data	5.2	5.3	5.4	5.4	5.5	5.6	5.6

Sample median = 5.4

$ x_i - X $	0.2	0.1	0.0	0.0	0.1	0.2	0.2
Ordered Difference	0.0	0.0	0.1	0.1	0.2	0.2	0.2

Ordered Difference 0.0 0.0 0.1 0.1 0.2 0.2 0.2

Therefore  $MAD = 0.1$

MAD is then scaled by a factor of 1.483 to make it equivalent to a normal deviation (MADE).

Hence  $MADE = 1.483 \times MAD = 0.1483$

If MADE is equal to zero SMAD should be calculated:

$SMAD = \text{mean} ( |x_i - X| \text{ } i = 1,2,\dots,n) \times 1.2531$

The Robust Standard Deviation may be used as the Standard Deviation for Proficiency

Assessment (SDPA) for calculation of z-scores. Other statistical methods for the calculation of

Robust estimators are available.

The Robust Standard Deviation may be used as the Standard Deviation for Proficiency Assessment (SDPA) for calculation of z-scores. Other statistical methods for the calculation of robust estimators are available.

## Removal of errors and blunders

Although robust estimators are used in order to minimise the influence of outlying results, extreme results or results that are identifiably invalid should not be included in the statistical analysis of the data. For example, these may be results caused by calculation errors or the use of incorrect units. However, such results can be difficult to identify by the PT organiser. For this reason, the robust mean and standard deviation will be calculated as above, but those results that are out of the range of the assigned value  $\pm 5 \times$  SDPA will be excluded and the robust mean and standard deviation will then be recalculated. These recalculated values will be used for the statistical analysis. All results, including excluded results, will be given performance scores.

## Estimated Standard Uncertainty of the assigned value

The assigned value ( $x_{pt}$ ) has a standard uncertainty ( $u(x_{pt})$ ) that depends upon the method used to derive the assigned value. When the assigned value is determined by the consensus of participants' results, the estimated standard uncertainty of the assigned value can be calculated by;

$$u(x_{pt}) = 1.25 \times \text{Robust standard deviation} / \sqrt{n} \quad \text{where } n = \text{number of results}$$

When the assigned value is determined by formulation, the standard uncertainty is estimated by the combination of uncertainties of all sources of error, such as gravimetric and volumetric measurements.

If  $u(x_{pt})$  is  $\leq 0.3 \times$  SDPA, then the uncertainty of the assigned value can be considered negligible and need not be considered in the interpretation of results.

If  $u(x_{pt})$  is  $> 0.3 \times$  SDPA, then the uncertainty of the assigned value is not negligible in relation to the SDPA and so  $z'$  ( $z$  prime) scores, which include the uncertainty of the assigned value in their calculation, will be reported in place of  $z$  scores.

z' scores are calculated as follows:

$$z' = \frac{(x_i - x_{pt})}{\sqrt{\sigma_{pt}^2 + u(x_{pt})^2}}$$

Where

$x_{pt}$	=	the assigned value
$x_i$	=	participant result
$\sigma_{pt}$	=	standard deviation for proficiency assessment
$u(x_{pt})$	=	standard uncertainty of the assigned value $x_{pt}$

$$\text{Expanded SDPA} = \sqrt{\sigma_{pt}^2 + u(x_{pt})^2}$$

The magnitude of z' scores should be interpreted in the same way as z scores.

### Sample ranges around the CV Cut off

We are often asked for help with interpreting reports, especially for qualitative analytes when there is no consensus. If we have a sample with 40% of results reported as negative and 60% as positive, what is the "correct" result? It's fairly simple. We could always send out very high positives and totally vacant negatives – and reduce our hassle factor. But obviously we must send out samples with borderline concentrations, hassle or no hassle. The simple answer to this problem is that if this was a patient than 40% of labs would tell him that he was negative and 60% would tell him he was positive. It may not be a nice concept to accept but it's the stark truth that the analyte and sample don't know there's a cut-off! The mean plus and minus the 2 SD range can straddle the cut-off. So, what is the "correct" result? In this example, chosen because it's extreme, both positive and negative are acceptable results!

### Acceptable but outside 2 SDs

And how do you interpret the situation where you are beyond the so-called-magical +/- 2 SDs but the clinical CV gives you an A for acceptable?

The clinical CV, or Acceptable Range, was added to the reports some years ago because some of the SDs was becoming seriously tight. Now, that sounds good but it doesn't matter how many results are in a data base, or how close they all are, the fact remains that 5% of labs will be told they are outside +/- 2 SDs – even if the difference between their result and the mean is insignificant. By definition, +/- 2 SDs includes 95% of results!

So, if you get the above situation, look at the actual figures in the report. Work out how far you are from the mean and if it insignificant (and you should know what that means, but if not look at the ranges you use for internal QC and assess it that way) then write that down on your action sheet or wherever your SOPs tell you to log potential problems or non-conformances, and move along.

And in case you doubt this, read ISO 13528. It states that it is acceptable to set “SDs” at a value required for a specific task (Section 6.2.1), or at a value that corresponds to the level of performance that the coordinator and members of the scheme would wish labs to achieve (Section 6.3.1). The example given is an acceptable 10% performance limit for glucose (Section 7.1.2). This sounds remarkably like what we’re trying to achieve with our Acceptable Ranges! These Acceptable Ranges are an ISO-acceptable way of setting standards of performance!

## Inside 2 SDs and A for Acceptable BUT consistently biased

If you find your results always on one side of the mean, but always within 2 SDs and Acceptable, pause and look at the real results. If you are Acceptable, it means you are not sending out clinically poor results; and being inside 2 SDs means you are NOT in the poorest 5% of results.

This MIGHT be a case for doing nothing. It depends on the size of your bias, the validity of the data base, and which analyte it is.

## Important General Requirements for ISO 17043 covered in our reports

Information on the scheme co-ordinator and authorisation of the report can be found in the Quality Control section of the report.

No Activities of Thistle QA are subcontracted.

**Confidentiality:** Thistle QA will not reveal or discuss lab performances on any of our EQAs without the written permission of the laboratory concerned.

**Scheme design:** Each EQA has been designed according to the relevant SOP; document number MQP-002, which is available on request

**AdCom:** Each scheme has an appropriately qualified and experienced AdCom. AdCom member details are kept on file at Thistle QA. The function of these committees is to advise Thistle QA on scheme design if necessary, and to handle disputes between client laboratories and Thistle QA with respect to differences of opinion, for example on scoring for a particular microorganism or the appropriate Clinical CV for an analyte.

**Traceability:** Most EQAs use consensus to establish the target value or mean and thus traceability is not required.

Where the company supplying the material has target values or supplies CRMs, certificates will be available on request.

**Homogeneity and stability:** Preparation of the samples and homogeneity studies are not performed at Thistle QA since prepared homogeneous material is purchased from the suppliers. Stability data is available on request

**Copyright:** The reports produced by Thistle QA are for the internal use of each laboratory or group of laboratories and may not be published or utilised for promotional purposes in any way without the written permission of the Director.

**Issue:** The Report Issue date and Version can be found on the cover page of the report.

**Accreditation:** All Schemes/Analytes are accredited to SANAS unless stated

Pilot Studies are not SANAS accredited EQA programmes



Analytes not accredited will have a flag on the Report stating “not part of the accreditation scope”

**Headers:**

Each Report page besides the cover page will have the Lab ID, Laboratory Name, Scheme, Round ID and Description as well as the Issued date.

**Footers:**

Each Report page including the cover will contain the Issue Number, Page numbers and Issue Date.

The last page of your report can be determined by the page numbers

**Support:**

For Further details and assistance with understanding your report better, please feel free to email [Thistleqa@lgcgroup.com](mailto:Thistleqa@lgcgroup.com) and we will gladly assist you.