

Table of Contents

1	Scope	2
2	Distribution	2
3	Definitions	2
4	Results shall be fairly dealt with	2
4.1	On time results	2
4.2	Late/Revised results	2
4.3	Data assessed “as reported” by Participants	2
4.4	Only valid data will be used	3
4.5	Multiple Test Results From a Single Laboratory	3
4.6	Data is available	3
5	Assessment of Quantitative Data	3
5.1	Z scores	3
5.1.1	Sufficient numeric results must be submitted	4
5.1.2	Result Precision (Electrical PT Programs ONLY)	4
5.1.3	Assessment of Microbiological Test Results	4
5.2	Prescribed Values	6
5.3	Non-Numeric Results for Quantitative Tests	6
5.3.1	Assessment of “Greater than” (>) and “Less than” (<) values	7
5.3.2	Assessment of other non-numeric results	7
5.4	Presentation of Quantitative Results in PTP Reports	7
5.4.1	Quantitative results for Enumeration Tests	7
6	Assessment of Qualitative Data	12
6.1	Categorical / Positive/ Negative Results	12
6.2	Purely Qualitative Results	13
6.3	Presentation of Qualitative Results in PTP reports	14

1 Scope

This policy applies to all Proficiency Testing Programs (PTP) conducted by IFM Quality Services Pty Ltd.
 Notes in green are explanatory only.

2 Distribution

This document is publically available.

3 Definitions

PTP	Proficiency testing program
IQR	Inter-quartile range
NIQR	Normalized inter-quartile range
Quantitative data	Numerical data upon which calculations can be performed
Qualitative data	<p>Also called categorical variables or attribute variables. The values of a qualitative variable can be put into a limited number of categories or different groups. Qualitative data may or may not have some logical order. Examples include:</p> <ul style="list-style-type: none"> • Absent/Present (organism detection) • Positive/Negative (oxidase reaction) • Sensitive, intermediate and resistant (antibiotic sensitivity) • 25 V, 50 V, 75 V, 100 V, 125 V (tracking voltage)

4 Results shall be fairly dealt with

4.1 On time results

- Once the due date has passed, all submitted results are considered to be final and cannot be withdrawn or amended. *This event is considered to be equivalent to the provision of a test report to a customer.*
- Only results received by the due date will be used to determine the assessment criteria.
- Results received by the due date will be assessed against the consensus values derived from these results.

4.2 Late/Revised results

- Late results are those received after the due date. Revised results include additional information or results submitted after the due date.
- A “late fee” is applicable prior to the acceptance of late results by IFM.
- Will be assessed against the consensus values derived from on-time results.
- Cannot be accepted after the PTP report has been issued.
- Will be recorded under separate participant entries.
- Will be marked as “Late” or “Revised” and contain a reference to the original participant as appropriate.

4.3 Data assessed “as reported” by Participants

From time to time, submitted results may appear to be incorrectly entered, such as evidenced by an obvious typographical or calculation error. However, IFM must maintain objectivity and cannot make assumptions with respect to a participant’s intent. Therefore, IFM cannot change a participant’s result

submission. IFM may however provide an override result value in order to achieve a standardized format required for result comparison and analysis. At no time will the original result be changed or removed.

4.4 Only valid data will be used

- When the PTP has included an equipment audit as part of the same PTP, and the equipment audit has revealed deficiencies (or the audit has not been completed), the integrity of the equipment and therefore the validity of test results are in doubt. Even though the laboratory's equipment may have been rectified prior to testing the PTP samples, IFM cannot verify that this has occurred. Therefore, test results arising from laboratories with deficient or incomplete equipment audits are excluded from the data population used to determine the consensus values. All submitted results will be assessed against the consensus result values.
- When the participant submission includes formats or characters that cannot be clearly and unambiguously interpreted, IFM will not provide an assessment. Participants need to take care to submit data in the requested format.

4.5 Multiple Test Results From a Single Laboratory

IFM strongly encourages all laboratory analysts to participate in PTPs in order that laboratories can best meet other requirements of ISO/IEC 17025. As a result, IFM provides result entry for multiple participants per laboratory.

However, when a laboratory submits a large number of results for a test or tests, this can create a bias in the consensus values and subsequently impact the calculation of the assessment criteria. To ensure all laboratories have equal influence in the calculated statistics, for each test, only the first 2 participants' results from each laboratory are included in the data population undergoing statistical analysis. Results from participant 3 and above will be excluded from this population. In charts, these results are shown in the bars via a separate colour indicating "Excluded".

Results that are not included in the statistical calculations will still be assessed against the determined consensus values.

4.6 Data is available

Although full result data is infrequently detailed in reports, the raw test data and the assessments are available on request to the PTP Program Coordinator.

5 Assessment of Quantitative Data

5.1 Z scores

Where possible, consensus statistics are used to assign Z scores.

$$Z = \frac{R - M}{NIQR}$$

Where: R = participant result , M = median (from consensus) , NIQR = the normalized inter-quartile range (i.e., 0.7413*[quartile 3 – quartile 1] for each set of result data).

Note: The NIQR is less affected by erroneous data points than the conventional SD. Therefore the range of acceptable results is more realistic and more easily determined (see pp. 11-21 of the below tutorial).
http://www.ifmqs.com.au/images/tutorial_PTP_200407.pdf

5.1.1 Sufficient numeric results must be submitted.

Z scores will only be assigned when 8 or more numeric data points are submitted.

When there are fewer than 8 numeric results, wherever possible, consensus and homogeneity data are used to provide pass / fail assessments.

5.1.2 Result Precision (Electrical PT Programs ONLY)

IFM calculates the statistically acceptable high and low as $M+3*NIQR$ and $M-3*NIQR$ respectively (Where M = median and NIQR = the normalized inter-quartile range). To ensure that participants reporting results to different levels of precision are not unfairly assessed, IFM rounds the statistically acceptable range to the same number of decimal points as the reported result. See the examples in **Table 1** below:

Table 1: Assessment of results based on the precision implied by the result

Result	Statistically acceptable low	Statistically acceptable high	Rounded statistically acceptable low	Rounded statistically acceptable high	Outcome
1	1.1000	2.0000	1	2	PASS
1.0	1.1000	2.0000	1.1	2.0	FAIL
2	1.0000	1.9888	1	2	PASS
2.0	1.0000	1.9888	1.0	2.0	PASS
2.0	1.0000	1.9400	1.0	1.9	FAIL
2.000	1.0000	1.9888	1.000	1.989	FAIL

5.1.3 Assessment of Microbiological Test Results

5.1.3.1 Acceptance of 0 cfu / unit as a result.

While a laboratory may not detect any targets during testing, this does not mean that there are no targets present in the material from which the sample was taken. It may mean that there were no targets present in the portion tested or that there were targets present but they were not recovered. For this reason, it should be noted that 0 cfu/unit is not considered a technically valid result. IFM therefore discourages reporting 0. Instead, it is considered to be more technically relevant to report “<1”, “<10”, etc according to the detection limit.

5.1.3.2 Raw Data versus Logarithm Transformed Data

For microbiological counts, statistics are based on either raw data or logarithm-transformed data. This is dependent on the median result value:

- When the median result is less than 100 cfu / (unit), statistical analysis, assessments and the assignment of Z scores will be based on raw result data. Results of 0 cfu / unit are included in the statistically assessed data population. Results of 0 cfu / unit will also receive a Z score.
- When the median result is greater than or equal to 100 cfu / unit, statistical analysis, assessments and the assignment of Z scores will be based on logarithm-transformed result data. Results of 0 cfu / unit are not included in the statistically assessed data population. (The logarithm of 0 is impossible to calculate.) Results of 0 cfu / unit will receive a FAIL assessment.

IFM may choose to assess data based on raw results in cases where the expected counts are low, even when the median value is greater than 100 cfu/unit. In these cases, a result of 0 cfu/unit will be included in the data population to be statistically assessed, and will receive a Z score.

Note 1: Higher counts are normally determined by performing serial dilutions. These are usually 10 fold. The numbers reported for high counts are approximations. They are NOT discrete numbers. Microbiological results are logarithmically transformed as approximated results do not form normalised data curves of distribution. This normal distribution of data is required for calculation of Z scores.

Note 2: The log value can be easily converted back to approximate real numbers (in cfu) by calculator, or using a spreadsheet.

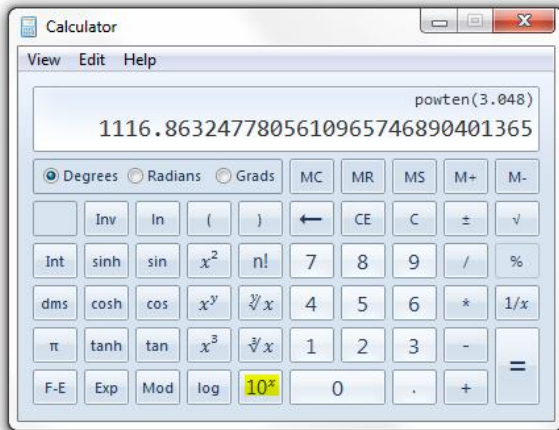
To convert a log number to a whole number (in cfu) in a calculator: is $10^{(\text{to the power of})}$ the log number (e.g. $10^5 = 100,000$ or $10^{3.0480} = 1117$)

Example: Coliform Count (Log values)

Low Value:	3.0480	= $10^{3.0480}$	= 1116.9
High Value	4.6905	= $10^{4.6905}$	= 49034.3
Median value	3.8692	= $10^{3.8692}$	= 7399.5

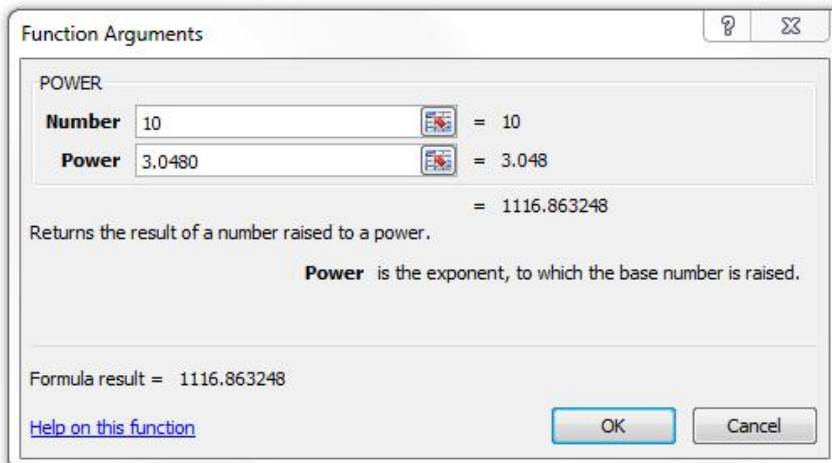
Desktop calculator '10^x power' function (Low Value)

Enter the power (exponent) value and click the 10^x function button



Excel Power function (Low Value)

Formula: $f_x = \text{POWER}(10, 3.0480)$



5.2 Prescribed Values

When assessment of results can be performed based on criteria from an indisputable source, results that fall within the acceptable range will receive pass assessments while those outside of the range will receive fail assessment.

For example, this method of assessment may be used if a standard has a requirement regarding the ambient temperature and the laboratory is required to report the ambient temperature.

5.3 Non-Numeric Results for Quantitative Tests

IFM encourages laboratories to think about the validity of submitted test results, both to customers and in PTPs. The intent of ISO/IEC 17025 makes it unacceptable to report results that imply greater precision or accuracy than can be technically justified. When the obtained value is beyond the limits of reliable measurements for the equipment or method in use, the analyst needs to qualify their result to convey an appropriate level of precision in the test report. Reporting of non-numeric data or qualified results allows the analyst to provide useful information to their customers (and the PTP provider) without compromising the intent of ISO/IEC 17025.

Non-numeric values arising from such circumstances are frequently “greater than” or “less than” results.

Examples:

1. A microbiological test method considers the acceptable range of bacterial colonies to be between 25 and 250 on an agar plate. 320 colonies are counted on the highest dilution originally prepared by the analyst. This is beyond the range defined by the method. The laboratory has not prepared sufficient sample dilutions to determine the number of colonies by the method and must indicate this fact in their reported results. They may report “>250” (multiplied by the dilution factor) or provide the count they actually obtained together with some indication the countable range was exceeded, (e.g., “320 estimated”).
2. In an electrical test, the voltage meter used for a determination has been calibrated over a specific range, but the measured value is outside that range. The laboratory cannot provide a valid result without conducting additional verifications. Alternatively, they may qualify their result to indicate it is outside the calibrated range.
3. An analyst may have noticed interference of non-target entities affecting their measured results, e.g. ambient voltage (noise), interfering chemicals or microbes that enhance or hide the true measurement, or cause cross-reactions. Such circumstances would require the analyst to report their result without implying a level of precision or accuracy beyond that justifiable by the analyst.

Statistical evaluation of non-numeric results submitted for enumeration tests is not possible. In such cases, alternate forms of assessment may be performed.

Generally, PASS or FAIL are allocated without calculating Z scores.

5.3.1 Assessment of “Greater than” (>) and “Less than” (<) values

The limit implied by the reported greater than or less than value is compared to the statistically acceptable ranges and assessments are allocated accordingly. See the examples in **Table 2** below:

Table 2: Assessment of Greater than and Less than values

Reported result	Statistically accepted range	Allocated assessment	Comment / Explanation
<50	25 - 100	PASS	Values within the statistically accepted range include results that are less than 50.
<50	60 - 160	FAIL	50 is smaller than the lowest statistically acceptable value
<50	10 - 40	PASS	All values within the statistically acceptable range are less than 50.
<50	50 - 100	FAIL	While exactly on the limit, a result less than 50 is below the statistically acceptable range.
>50	25 - 100	PASS	Values within the statistically accepted range include results that are greater than 50.
>50	60 - 160	NOT ASSESSED	When the maximum detection limit is less than the acceptable low, the participant is not assessed by IFM. Instead, the laboratory should self assess by considering if the result would meet a customer’s needs.
>50	10 - 40	FAIL	All values within the statistically acceptable range are less than 50.
>50	50 - 100	PASS	While exactly on the limit, a result greater than 50 is within the statistically acceptable range.

Exception: For microbiological tests, the highest “less than value” that will be accepted is 100 cfu / unit. Values, such as <1000 cfu / unit will not be assessed.

5.3.2 Assessment of other non-numeric results

- Responses reflecting the assigned value will receive a PASS assessment.
- Responses not reflecting the assigned value will receive a FAIL assessment.
- In cases where the statistically acceptable range extends beyond the usual lower limits of detection to 0 cfu / unit, a result of not detected, absent or similar, will receive a PASS assessment.
- For microbiological tests only: In cases where the target organism is absent, a result of “detected”, “present” (or similar), will receive a FAIL assessment¹

5.4 Presentation of Quantitative Results in PTP Reports

5.4.1 Quantitative results for Enumeration Tests

5.4.1.1 Case 1: Target analyte is absent

- Results are represented using a stacked frequency histogram (See **Chart 1**)

¹ When more than 10% of laboratories report an unexpected result of Present, if not already undertaken, additional testing will be performed by IFM using more sensitive methods. Assessment of the affected participants will then take these test results into consideration.

- When participants incorrectly report a numerical result, (for example in a microbiological test where the target is absent), the result is counted as a “False Positive”.
- When participants correctly report a “0” or “less than” result, the result is counted as “Below Detection Limit”

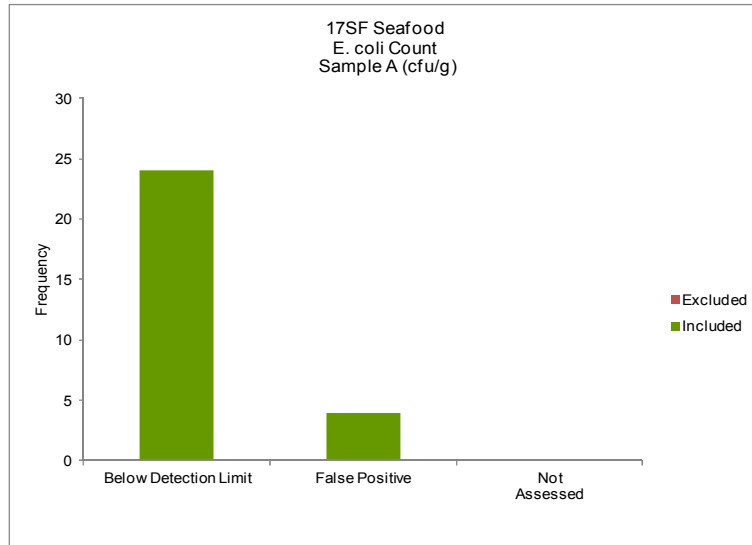


Chart 1: Example of Chart presentation where the target analyte is absent.

5.4.1.2 Case 2: Target analyte is present at very low levels

- Where participants have reported results that are below a particular detection limit and have been favourably assessed, these will be counted as Detected < x /unit where X is the level of detection. See **Chart 2**.

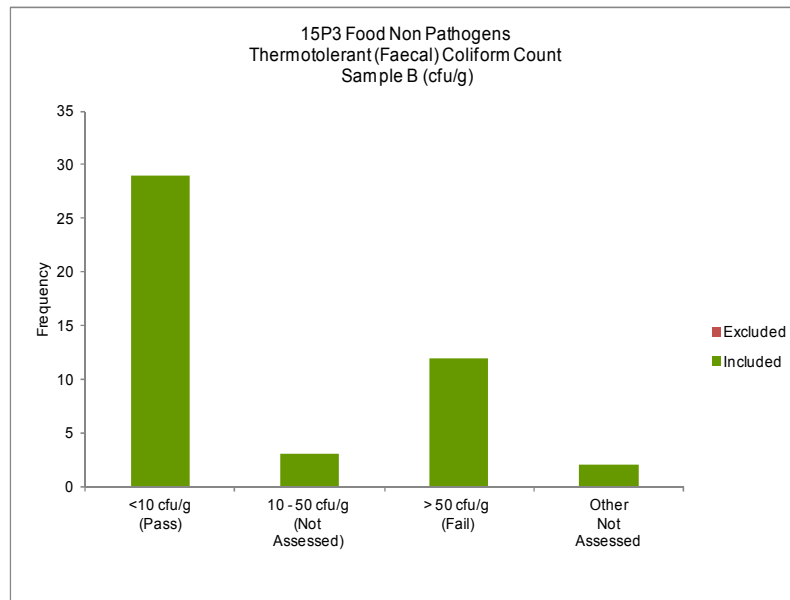


Chart 2: Example of Chart presentation where the target analyte is present at low levels.

5.4.1.3 Case 3: Target analyte is present at quantifiable levels

- Results are presented via a stacked frequency histogram. (See **Chart 3** and **Chart 4**.)

Note: All results submitted are displayed in reports.

- The X-axis is divided into 5 sections called “bins”.
 1. Below detection limit
 - Each column shows the number of results reported below the detection limit within the range of the bin.
 2. Numeric results
 - Numeric results are displayed in a frequency histogram.
 3. Above detection limit
 - Each column shows the number of results reported above the detection limit within the range of the bin.
 4. Detection
 - Section 4 displays results of present or absent, or similar.
 5. Other
 - Section 5 shows results that cannot be categorised into the above.

Note: The first bin represents all results up to and including the value of the label. Each subsequent bin represents values greater than the previous bin up to and including the value of the label. There may be Above/Below detection limit or ‘other’ columns. When present, these bins show the number of results reported within the range of the bin.

- The Y (vertical) axis indicates the number (or “frequency”) of results in each bar. Results that are included in the calculation of the median and NIQR are represented by the blue portion of the bar.
- Numeric results that are excluded from the calculation of the median and NIQR are represented by the red portion of the bar.
- Results that are not numeric and therefore cannot be included in the calculation of the median and NIQR are represented by the green bar.

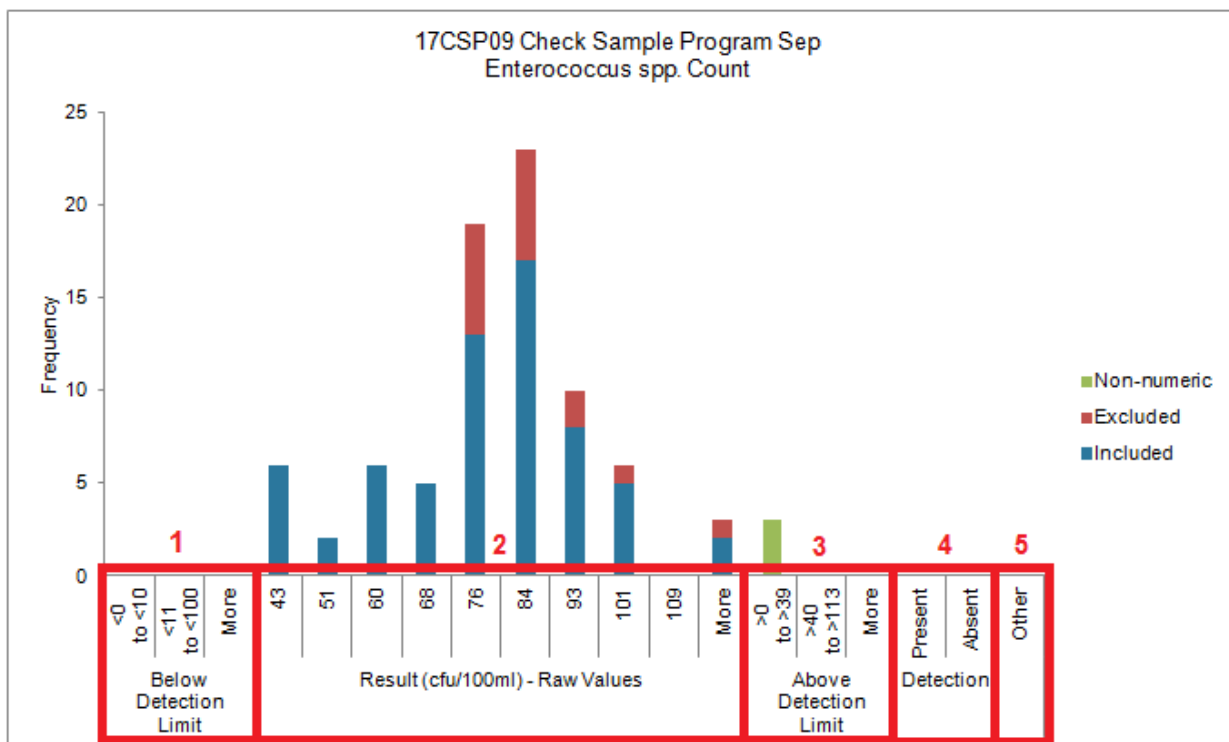


Chart 3: Example of Chart presentation where the target analyte is present at quantifiable levels.

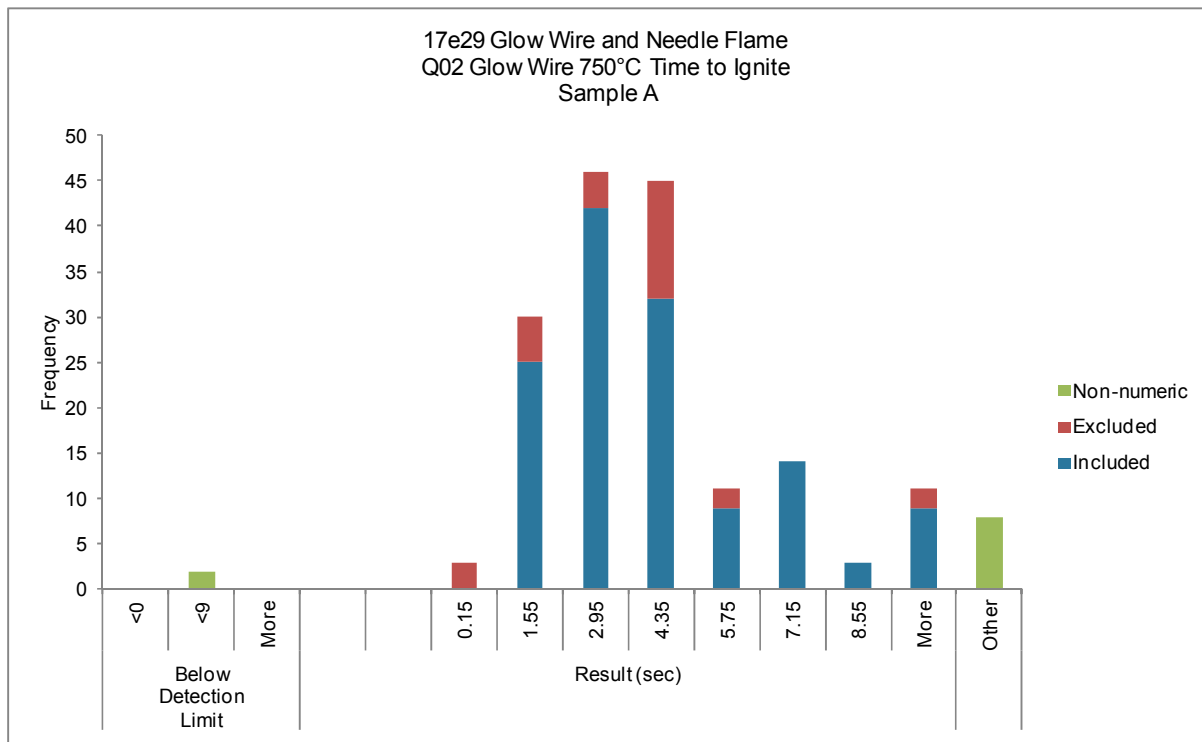


Chart 4: Example of Chart presentation where the target result is quantifiable

Note: IFM ensures that the data are normally distributed before being used for statistical analysis. The charts in our report may sometimes appear not to be normally distributed; this is frequently due to the truncation of results for inclusion in a chart. The results up to the first value are included in the first bar and all that are higher than the range under consideration are included in the ‘more’ bar.

5.4.1.3.1 Statistical Tables

Statistical tables provide information such as the number of results analysed, the median, the statistically acceptable range, etc. (See **Table 3**.)

Note 1: The statistical summary table is presented so that participants can verify their own Z scores and compare their result/s against the acceptable range of values and spread of participant results. The statistical tables are presented in either raw or log values.

- All results are first placed in numerical order. Median is the middle score
- Q1 is the value under which the first 25% of participants reported as results
- Q3 is the value under which the first 75% of participants reported as their result
- The interquartile range (IQR) is the range of results reported by the middle 50% of participants
- The normalized interquartile range (NIQR) is the interquartile range multiplied by 0.7413. NB. This factor is applied to a normal distribution of results
- Acceptable high/ low value is the median result + or – (3 x NIQR).

Table 3: Example of Statistical Summary Presentation

16P3FP Food Pathogens Clostridium perfringens Count (vegetative) Sample C (cfu/g) - Log Values	
Number of Results Received	38
Number of Results Analysed	32
Quartile 1	3.6968
Quartile 3	4.0202
IQR	0.3234
NIQR	0.2397
Median	3.8543
Acceptable High: Median + (3*NIQR)	4.5734
Acceptable Low: Median - (3*NIQR)	3.1352
Statistically Acceptable Range	1.4382

Note 2: Sometimes it is necessary for a participant to assess their own results (e.g. if results were not submitted on time, or in conducting corrective actions). The acceptable high and low values are found in the summary tables previously described, and highlighted in examples below. If the participant's result falls between these values it is considered to be acceptable.

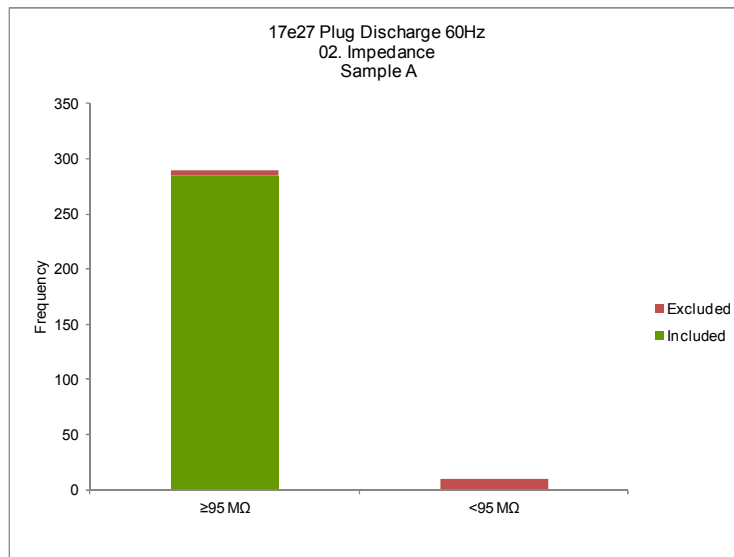
18e37 Working Voltage 50Hz 240V ac Q04.1 Volts Pos1-2 II/Centre rms (Volts)	19P5 Pharmaceutical per 1.0g Plate Count Sample B (cfu/g) - Log Values	19MIS06 Meat Scheme Jun Coliform Count (cfu/g) - Raw Values
Number of Results Received	66	43
Number of Results Analysed	58	34
Quartile 1	4.1959	6.0000
Quartile 3	4.5563	10.0000
IQR	0.3604	4.0000
NIQR	0.2672	2.9652
Median	4.4472	7.5000
Acceptable High: Median + (3*NIQR)	5.2487	16.3956
Acceptable Low: Median - (3*NIQR)	3.6457	-1.3956
Statistically Acceptable Range	1.6030	17.7912

In order to calculate Z scores for results (based on program consensus data), participants may use the spreadsheet below:

[http://www.ifmqs.com.au/proficiency/references/Z-score Calculator_V3.xls](http://www.ifmqs.com.au/proficiency/references/Z-score%20Calculator_V3.xls)

5.4.1.4 Case 4: Assessment is based on indisputable criteria

- The discussion will detail the assessment criteria
- The number of participants reporting results for each category described in the discussion will be shown in the chart



6 Assessment of Qualitative Data

6.1 Categorical / Positive/ Negative Results

In the examples below, “POSITIVE” refers to all forms of results that imply a positive test outcome. For example: “tracking”, “ignition”, “growth”, “detection”, “pass”, “yes” are all results equivalent to “POSITIVE” in the below table. Conversely, “no tracking”, “no ignition”, “no growth”, “not detected”, “fail”, “no” are results that are considered “NEGATIVE” in the below table.

Refer to **Table 4**:

Table 4: Assessment of Positive and Negative Qualitative Results

Expected Outcome	75 % consensus	Participant reported	Assessment
POSITIVE	POSITIVE	POSITIVE	PASS
		NEGATIVE	FAIL
POSITIVE	NEGATIVE	POSITIVE	PASS
		NEGATIVE	PASS
POSITIVE	No Consensus	POSITIVE	PASS
		NEGATIVE	Not assessed
POSITIVE	POSITIVE	Numerical result	PASS ²
NEGATIVE	NEGATIVE	POSITIVE	FAIL
		NEGATIVE	PASS
NEGATIVE	POSITIVE	POSITIVE	PASS
		NEGATIVE	PASS
NEGATIVE	No Consensus	POSITIVE	Not Assessed
		NEGATIVE	PASS
NEGATIVE	NEGATIVE	Numerical result	FAIL ³

Exception: where the expected outcome is derived from an indisputable source, such as a testing standard or well established characteristics of a microorganism, participants reporting results contrary to the expected outcome will be awarded a fail assessment regardless of the consensus.

² The result is treated as “positive”.

³ The result is treated as “positive”.

6.2 Purely Qualitative Results

For purely qualitative results, such as identification of microbes, the result is assessed against the level of confidence implied by the submitted result. Refer to Table 5 below, which provides an example.

Table 5: Assessment of “Pure” Qualitative Results (Example)

Organism provided	Result Reported	Result Assessment	Comment / Explanation
<i>Legionella pneumophila</i> serogroup 1	Unidentified Gram negative rod	PASS	Though incomplete, the reported information provided is correct. (It is the laboratory’s responsibility to determine whether their identification is adequate.)
	<i>Legionella</i> species	PASS	The reported genus is correct. No other claim has been made by the participant. ⁴
	<i>Legionella pneumophila</i>	PASS	The reported genus and species are correct.
	<i>Legionella gormanii</i>	FAIL	The reported species name is incorrect
	<i>Legionella pneumophila</i> SG1	PASS	The reported genus, species and serogroup are correct
	<i>Legionella pneumophila</i> SG4	FAIL	Although the correct genus and species have been reported, the serogroup is incorrect.

⁴ if identification to species level was requested this result would not be assessed as insufficient information was provided

6.3 Presentation of Qualitative Results in PTP reports

- Results are presented via a stacked frequency histograms. Refer to **Chart 5** and **Chart 6**.
- Results are shown in the bars corresponding to the assessed categories.
- Where relevant, other values are accumulated under “Other”.
- Those results included in the determination of the consensus value, are represented by the green portion of each bar.
- Those results excluded from the determination of the consensus are represented by the red portion of each bar.

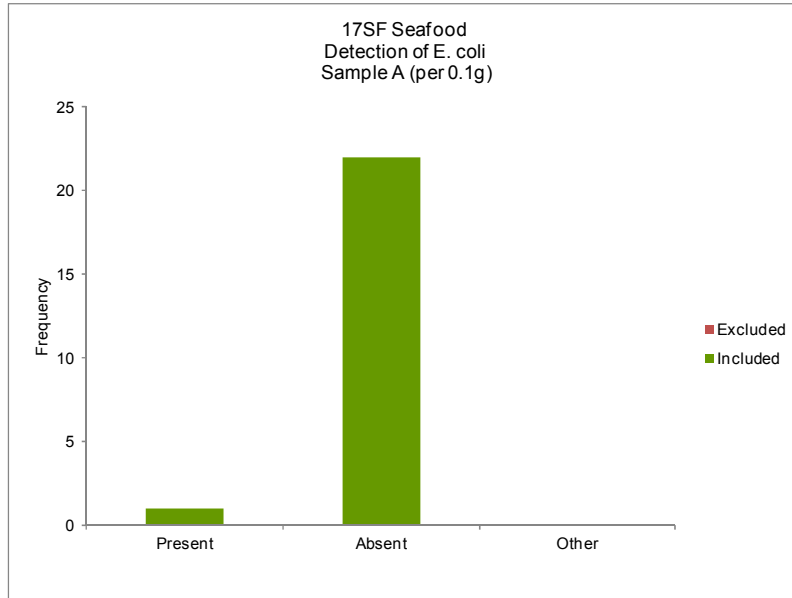


Chart 5: Example of presentation of Qualitative Results

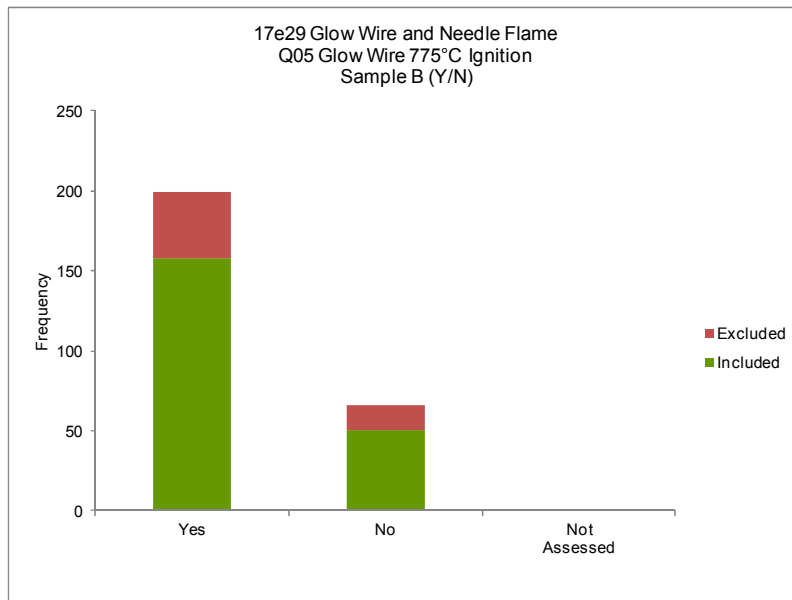


Chart 6: Example of presentation of Qualitative Results