INTRODUCTION

The examples provided are intended to demonstrate ways to implement the A2LA policies for the evaluation of measurement uncertainty (MU) for methods that use counting for determining the number of colony forming units (CFU) or most probable number (MPN) in a test sample. These examples apply for all quantitative microbiological methods. To the extent possible they use conventional concepts for evaluating measurement uncertainty, but they are not the only possible methods for determining measurement uncertainty of counting methods.

This document is meant to provide a practical approach to meeting the requirements for evaluating measurement uncertainty defined in ISO/IEC 17025:2017 for quantitative microbiology testing. These examples are based on 20 or 30 data points, but larger datasets will produce more reliable estimates and smaller data sets may be used with caution. The coverage factor obtained from the Student t-tables should be used to estimate expanded uncertainty for smaller datasets.

The premise behind these examples is that the laboratory has already verified the standard method they are using or validated any non-standard or modified method in use, including determining acceptability of specific matrices.

The VIM definition of precision includes the phrase “under specified conditions”. These conditions are uncertainty components. So, there is a requirement to identify the conditions varied and not varied. For example, when determining intralaboratory reproducibility, the conditions of different days and different analysts may be varied. Different instruments may not be included if the laboratory only has one instrument at the time. If a new or additional instrument is obtained, the laboratory knows that the impact of a different instrument was not included in the intralaboratory reproducibility determination. If instrumentation is a significant contributor to the uncertainty, the laboratory must estimate the uncertainty contributed by instrumentation before the additional instrument is placed into use.

In order to evaluate the measurement uncertainty, the potential uncertainty components (conditions) are identified. Then data are obtained to estimate the uncertainty components. It may not be practical or possible to estimate all uncertainty components. The record for the evaluation of measurement uncertainty needs to identify the uncertainty components that could not be estimated. This is more easily done if the records for measurement uncertainty clearly identify the conditions varied (uncertainty components estimated). To assist in identifying the potential uncertainty components Table 5 is provided.

For each of these examples, the $\log_{10}$ result was used instead of the natural log for ease of calculation. Conversion to $\log_{10}$ is necessary because of the nature of the measurand (Colony Forming Units), which reproduces exponentially and the dilution scheme, which is determined by serial dilution.

 A2LA confidential document. A2LA Copyright.

PROCESS TO EVALUATE MEASUREMENT UNCERTAINTY:

1. Identify the target analyte(s) including units (measurand), examples include:
    a. Coliforms (MPN/g)
    b. Staphylococcus aureus (CFU/g)
    c. Mesophilic aerobic bacteria (CFU/mL)
    d. Total fungal count (CFU/swab)
    e. Generic Escherichia coli (MPN/mL)
2. Identify potential uncertainty components, examples include (but not limited to):
    a. Analyst (collection of test portions and preparation of samples)
    b. Incubation temperature
    c. Incubation time
    d. Dilution blank preparation
    e. Media preparation
    f. Media temperature
    g. Media pH
    h. Balance
    i. Pipet/pipettor
    j. Standard uncertainty of spike concentration
    k. Standard uncertainty of a property value for a Certified Reference Material (if used)

    Note: Table 5 lists all these potential uncertainty components in an organized manner so the laboratory can illustrate which components are relevant to their method.
3. Identify best model for evaluating measurement uncertainty based upon method, available data, needs of the customer, and use of the measurement uncertainty. Examples of models include (but not limited to):
    a. Technical Uncertainty (example 1): Technical uncertainty is the operational variability associated with the technical steps of the method. This model is used when control samples run through all steps of the process and the same target value is used with each run. Another model might be needed if MU is used for statements of conformity.
    b. Recovery Replicates (example 2): Recovery replicates are the comparison of the recovery of an organism with and without matrix. This model is used when laboratory control sample is not the same concentration level with each run.
    c. Intralaboratory Reproducibility Replicates, also known as Intermediate Precision Replicates, (example 3): Intralaboratory Reproducibility Replicates are either control samples set up in replicate (example 3a) or actual samples received in duplicate and analyzed (example 3b), with each analytical run. This model might be used when statements of conformity are needed.
    d. Validation Study MU (example 4): Validation study MU uses the technical data from the method validation in order to calculate the uncertainty of the method validated. This model is used when validation data are available, and the resulting MU meets the needs of the laboratory.
4. Calculate the expanded uncertainty.
    a. By convention the results are always rounded down for the value at the lower end of the range and always rounded up for the value at the upper end. In this way the minimal 95% coverage is preserved.

EXAMPLE 1: TECHNICAL UNCERTAINTY USING LABORATORY CONTROL SAMPLE

For this example, the laboratory quality control material is created by spiking a known quantity into a matrix such as nonfat dry milk, tissue, or other relevant matrix. The matrix is representative of the laboratory samples the laboratory usually tests. This is the Laboratory Control Sample (LCS).

This approach is defined in P103b, Annex: Policy on Estimating Measurement Uncertainty for Life Sciences Testing Laboratories for Category III methods. Laboratory Control Sample results may be used to evaluate Measurement Uncertainty, provided the samples are an appropriate matrix and concentration.

When the LCS has been through all method steps and includes significant uncertainty components, then the laboratory can use the standard deviation (SD) as an estimate of combined standard uncertainty. A Relative Standard Deviation (RSD) or Coefficient of Variation (CV) may also be used.

It is recommended that 20 or more individual LCS data points be obtained to estimate SD. The estimate of expanded uncertainty is then calculated using the formula:

Expanded Uncertainty for a Defined Method (LCS) = k x SD

where k (the coverage factor) equals 2 (for 95% confidence).

Note: in this example the SD includes the components identified for the combined standard deviation (SDC) (described in section "Additional Uncertainty Components" below).

If fewer than 20 LCS results are available, the coverage factor should be the appropriate t statistic for 95% confidence for the associated number of degrees of freedom (for example 5= 2.57, 10= 2.23, 20= 2.09, 30 = 2.04 and 60 = 2.00).

Table 1: Laboratory Control Data with Same Target Values (e.g., 100 CFU)

Raw Data (actual CFU recovered)	Log10 Value
131	2.1173
69	1.8388
45	1.6532
40	1.6021
31	1.4914
33	1.5185
31	1.4914
37	1.5682
186	2.2695
218	2.3385
200	2.3010
39	1.5911
217	2.3365
119	2.0755
28	1.4472
106	2.0253
107	2.0294
45	1.6532
98	1.9912
240	2.3802

The act of spiking the inoculum adds an uncertainty component caused by the uncertainty in creating the spike and in assigning a value to the spike. In microbiology, the variability amongst spikes is usually the significant uncertainty component, and this uncertainty will be captured in the recovery replicate data over time. If this is not the case for a laboratory's method, the additional uncertainty component of the spike value will need to be estimated.

Evaluating Uncertainty using the Standard Deviation:

Step 1. Transform the CFU values (column 1) to $\log_{10}$ value (column 2).

Step 2. Calculate the standard deviation of the $\log_{10}$ values. The SD of this data set is 0.3348.
Combine standard deviations if needed (In this example, there is no additional component of uncertainty and no bias adjustment. Only the SD of the data set is used).

Step 3. For reporting purposes, apply the coverage factor to the SD to obtain the expanded uncertainty. For 95% coverage we use a coverage factor of k=2. The expanded uncertainty in this example is thus 0.6696. If using the student t-tables, with n=20 (19 degrees of freedom), the coverage factor k would be 2.09, which would provide an expanded uncertainty of 0.6998.

Step 4. To calculate the measurement uncertainty as expanded uncertainty for any subsequent laboratory result using SD x k, the result is first converted to $\log_{10}$ value, and the expanded uncertainty of 0.6696 is added and subtracted from the log value.

Step 5. To estimate the MU of a sample result, calculate the uncertainty interval for the logarithmic value, by adding and subtracting the expanded uncertainty from the sample $\log_{10}$ value; then convert the log value for the sample measurement back to CFU for the reported result. This is accomplished by taking the anti-log of each of the endpoints of the interval (anti-log of $x = 10^x$).

For example, estimating the uncertainty using SD x k for a result of 150 CFU: 150 in $\log_{10} = 2.1761$. Adding and subtracting 0.6696 from 2.1761 gives an interval from 1.5065 to 2.8457; transforming back to counts: 101.5065 =32.10, and 102.8457 =700.97. Therefore, the uncertainty interval is 32 to 701 CFU.

Evaluating Uncertainty using the Relative Standard Deviation (RSD):

Step 1. Transform the CFU values (column 1) to $\log_{10}$ (column 2).

Step 2. Calculate the mean and standard deviation, SD, of the $\log_{10}$ values. The SD of this data set is 0.3348. The mean of this dataset $\log_{10}$ is 1.8860, or 77 CFU. This could also be expressed as a percentage standard deviation (or relative standard deviation), in log units: 0.3348 divided by 1.8860 = 0.1775, or 17.8%; an expanded relative uncertainty is 2 x 0.1775 = 0.3550, or 35.5%.

Step 3. To estimate the uncertainty for any subsequent laboratory result using RSD x k, the result is first converted to the $\log_{10}$, multiplied by 0.355 and then this expanded uncertainty is added and subtracted from the log value**.**

Step 4. To estimate the MU of a sample result, calculate the uncertainty interval for the logarithmic value, by adding and subtracting the expanded uncertainty from the sample $\log_{10}$ value; then convert the log value for the sample measurement back to CFU for the reported result. This is accomplished by taking the anti-log of each of the endpoints of the interval (anti-log of $x = 10^x$).

For example, if estimating the uncertainty using RSD x k for a result of 150 CFU: 150 in $\log_{10} = 2.1761$ and the expanded uncertainty in log counts is 2.1761 x 0.355 = 0.7725. Adding and subtracting from 2.1761 gives an interval from 1.4036 to 2.9486; transforming back to counts: 101.4036 =25.33, and 102.9486 = 888.41. Therefore, the uncertainty interval is 25 to 889 CFU.

Note: When using relative standard deviation of data sets, the uncertainty value needs to be multiplied by the test result it is being applied to, as the value is being compared and relative to the data produced. When using standard deviation of data sets, the uncertainty value does not need to be multiplied by the test result it is being applied to. The expectation is that the uncertainty is not relative to the concentration of the sample, and the uncertainty would be the same across all concentrations for that method.

Additional Uncertainty Components

If the QC material has not been run through all method steps, then the laboratory should incorporate any appropriate additional components or conditions in the uncertainty calculations. The standard uncertainty for these additional conditions is estimated (SDA for standard deviation additional). Examples of some additional conditions are those uncertainty components from sub-sampling, taking an aliquot, or sample preparation. The additional components should be combined with SD using the root sum square (RSS) method.

If a method has a known consistent bias that is inherent to the method (e.g., low recovery on difficult analytes) the bias must not be added to the uncertainty calculations. The bias shall, however, be clearly stated and recorded along with the uncertainty estimate. If a bias adjustment is made prior to reporting a result (e.g., adjusting for recovery on a sample that is spiked with a known amount of substance), then an additional component of uncertainty is introduced and must be included in the uncertainty estimate (SDB for standard deviation from bias adjustment). This component can also be combined using the RSS method. However, if LCS data routinely include adjustments for recovery, then the uncertainty from the adjustment is already included in SD and does not need to be added again.

The formula for the RSS method to combine standard uncertainties is:

$SDC = \sqrt{(SD^2+SDA^2+SDB^2)}$

SDC is the combined standard uncertainty

Note that for the data in example 1, SDA and SDB are both zero.

EXAMPLE 2: RECOVERY REPLICATES FOR LABORATORY CONTROL MATERIALS

Due to the behavior of organisms and their interaction with the matrix, it is expected that recovery is reasonably constant for a particular organism in a given matrix; hence, the uncertainty in microbiological measurements can be estimated by looking at recovery over time. In microbiology it is well documented that many organisms behave consistently, so this expectation is often met. Recovery differences over time should reflect the various components of uncertainty, including those identified in Table 5.

In this example (Table 2), the same amount of inoculum is plated both without matrix and with the matrix of interest. The difference between the counts from the plate without matrix and the plate with matrix samples is a measure of recovery of the organism, usually expressed as a percentage of the CFU count in the inoculated sample. The 20 replicates were of greatly different levels (as can be expected in some situations) but the recovery is reasonably constant. For this model, the matrix used must be representative of those samples where an estimate of uncertainty is needed.

Table 2: Recovery Replicates

CFU Inoculated	Log10 Value	CFU recovered in Spike	Log10 Value	% Recovery of Log Values
30,000	4.4771	20,000	4.3010	96.1
17,000	4.2304	12,000	4.0792	96.4
36,000	4.5563	49,000	4.6902	102.9
150	2.1761	90	1.9542	89.8
2,400	3.3802	1,300	3.1139	92.1
43,000	4.6335	32,000	4.5051	97.2
100	2.0000	98	1.9912	99.6
42,000	4.6232	31,000	4.4914	97.1
19,000	4.2788	12,000	4.0792	95.3
100	2.0000	120	2.0792	104.0
580,000	5.7634	410,000	5.6128	97.4
2,500	3.3979	2,000	3.3010	97.1
1,100	3.0414	930	2.9685	97.6
18,000	4.2553	12,000	4.0792	95.9

2,000	3.3010	1,900	3.2788	99.3
1,700	3.2304	2,100	3.3222	102.8
2,100	3.3222	1,700	3.2304	97.2
150	2.1761	100	2.0000	91.9
2,000	3.3010	1,600	3.2041	97.1
150	2.1761	110	2.0414	93.8

Step 1. Transform the CFU values (columns 1, 3) to $\log_{10}$ (columns 2, 4).

Step 2. Calculate the % recovery of the $\log_{10}$ values by dividing column 4 by column 2 and multiplying by 100 (column 5).

Step 3. Calculate the mean and standard deviation (SD) of the % recovery of the $\log_{10}$ values. The mean recovery is 97.0% and the SD of % recovery is 3.6%. The SD is an estimate of the combined standard uncertainty, which can be used as a relative uncertainty.

Step 4. For reporting purposes, apply the coverage factor to the SD to obtain the expanded uncertainty. For 95% coverage we use a coverage factor of k=2. The expanded uncertainty in this example is 7.2%. If using the student t-tables, with n=20 (19 degrees of freedom), the coverage factor k would be 2.09, which would provide an expanded uncertainty of 7.5%.

Step 5. Because the recovery is expressed as a percentage, when calculating the expanded uncertainty for a sample, this percentage needs to be multiplied by the $\log_{10}$ value in order to estimate the uncertainty in log units.

For example, for a result of 150 CFU: 150 in $\log_{10}$ = 2.1761 and the expanded uncertainty in log counts is 2.1761 x 0.072 = 0.1567. Adding and subtracting from 2.1761 gives an interval from 2.0194 to 2.3328, transforming back to counts: 104.57 to 215.18. Therefore, the uncertainty interval is 104 to 216 CFU.

Note: In this example, recovery is close to 100% and therefore there is no indication of significant bias. If average recovery were much lower (e.g., 80%) then there would be an indication of consistent bias that should be investigated. However, it is not current conventional practice to correct for bias in quantitative microbiology results.

EXAMPLE 3: INTRALABORATORY REPRODUCIBILITY REPLICATES FOR TEST SAMPLE OR LABORATORY CONTROL SAMPLES

EXAMPLE 3A: Samples usually do not contain CFU. The majority of test results are <1 or <10 CFU.

This procedure illustrates the use of intralaboratory reproducibility replicates to estimate uncertainty for the same type of sample matrix analyzed. This technique captures various sources of uncertainty that can affect routine customer samples, by having "replicates" be produced independently under as many different conditions as possible that normally occur in a laboratory.

The example data given in Table 3 are results from control samples that are analyzed through all steps of the test method and were set up in duplicate on different days, by different analysts, using different equipment (e.g., balances, pipettors) and as possible, using different batches of media/reagents. Since microbiology samples are often not stable over time, appropriate care must be taken if replicates are analyzed on different days. The control samples are made by spiking a matrix material. A requirement for this approach is that the matrix used to generate the data in Table 3 is representative of the samples analyzed by the laboratory and therefore this model may be used to evaluate the measurement uncertainty for any sample. This is useful for food or feed production laboratories where the same types of samples are analyzed for products to determine conformance to product specifications.

Table 3: Replicates Generated Under Reproducibility Conditions

Raw Data (actual CFU recovered) – First Replicate	Log_{10} Value	Raw Data (actual CFU recovered) – Second Replicate	Log_{10} Value	Difference between Replicates (Log_{10} Value)	Difference between Replicates Squared
131	2.1173	142	2.1523	-0.0350	0.00123
69	1.8388	90	1.9542	-0.1154	0.01332
45	1.6532	76	1.8808	-0.2276	0.05180
40	1.6021	55	1.7404	-01383	0.01913
31	1.4914	20	1.3010	0.1903	0.03623
33	1.5185	40	1.6021	-0.0835	0.00698
31	1.4914	62	1.7924	-0.3010	0.09062
37	1.5682	50	1.6990	-0.1308	0.01710
186	2.2695	167	2.2227	0.0468	0.00219
218	2.3385	258	2.4116	-0.0732	0.00535
200	2.3010	243	2.3856	-0.0846	0.00715
39	1.5911	54	1.7324	-0.1413	0.01997
217	2.3365	180	2.2553	0.0812	0.00659
119	2.0755	133	2.1239	-0.0483	0.00233
28	1.4472	46	1.6628	-0.2156	0.04648
106	2.0253	112	2.0492	-0.0239	0.00057
107	2.0294	89	1.9494	0.0800	0.00640
45	1.6532	62	1.7924	-0.1392	0.01937
98	1.9912	128	2.1072	-0.1160	0.01345
240	2.3802	220	2.3424	0.0378	0.00143

Step 1. Transform the raw data by taking the log_{10} of the data (column 2, 4).

Step 2. Calculate the difference between the transformed replicates (column 5).

Step 3. Square the differences between the transformed replicates (column 6).

Step 4. Add the differences together (column 6) and divide by 2n, where n = the total number of pairs of replicates (for this example n = 20) to get 0.00919.

Step 5. Take the square root of the result in step 4; this equals the pooled intralaboratory reproducibility standard deviation, which is 0.0959.

Step 6. To provide a higher range of values that cover what is likely to be observed, apply the coverage factor (k=2 for 95% coverage) to the intralaboratory reproducibility standard deviation to get the estimate of the expanded uncertainty of 0.192. (Note this is a log_{10} value).

Step 7. To calculate the uncertainty for any given result, the result is first converted to the log_{10} value and then the expanded uncertainty 0.192 is added and subtracted from the log result.

Step 8. To estimate the MU of a sample, convert the log value for the sample measurement back to base units (CFU) for the reported result. This is accomplished by taking the anti-log of the endpoints of the interval (anti-log of $x = 10^x$).

For example, for a sample result of 150 CFU: 150 in $\log_{10} = 2.176$, so the interval in $\log_{10}$ counts is 1.984 to 2.368; transforming to counts: $10^{1.984} = 96.4$, and $10^{2.368} = 233.3$. Therefore, the uncertainty interval is 96 to 234 CFU.

EXAMPLE 3B: Samples contain CFU and do not need to be spiked

When duplicate samples that contain CFU are available, spikes do not need to be made. The results for the samples themselves can be used. The same calculations are made.

The example data given in Table 4 are from laboratory samples of cannabis flower received in duplicate. Each laboratory sample is taken through the entire method. Since the analyte is incurred, no spiking is needed. There is no uncertainty added due to spiking. The range covered is from 100 to 1000 CFU as an average of the two values.

Table 4: Samples Received in Duplicate

Raw Data (actual CFU recovered) – First Replicate	Log_{10} Value	Raw Data (actual CFU recovered) – Second Replicate	Log_{10} Value	Difference between Replicates (Log_{10} Value)	Difference between Replicates Squared
160	2.2041	50	1.6990	0.5051	0.25518
90	1.9542	120	2.0792	-0.1249	0.01561
190	2.2788	30	1.4771	0.8016	0.64261
100	2.0000	130	2.1139	-0.1139	0.01298
160	2.2041	100	2.0000	0.2041	0.04166
80	1.9031	210	2.3222	-0.4191	0.17567
30	1.4771	260	2.4150	-0.9379	0.87957
180	2.2553	190	2.2788	-0.0235	0.00055
140	2.1461	270	2.4314	-0.2852	0.08136
120	2.0792	400	2.6021	-0.5229	0.27340
190	2.2788	490	2.6902	-0.4114	0.16928
430	2.6335	260	2.4150	0.2185	0.04774
180	2.2553	520	2.7160	-0.4607	0.21227
120	2.0792	580	2.7634	-0.6842	0.46819
450	2.6532	490	2.6902	-0.0370	0.00137
580	2.7634	370	2.5682	0.1952	0.03811
710	2.8513	250	2.3979	0.4533	0.20550
600	2.7782	460	2.6628	0.1154	0.01332
120	2.0792	940	2.9731	-0.8939	0.79914
590	2.7709	710	2.8513	-0.0804	0.00647

Raw Data (actual CFU recovered) – First Replicate	Log_{10} Value	Raw Data (actual CFU recovered) – Second Replicate	Log_{10} Value	Difference between Replicates (Log_{10} Value)	Difference between Replicates Squared
630	2.7993	680	2.8325	-0.0332	0.00110
700	2.8451	620	2.7924	0.0527	0.00278
820	2.9138	580	2.7634	0.1504	0.02262
580	2.7634	830	2.9191	-0.1557	0.02423
880	2.9445	540	2.7324	0.2121	0.04498
730	2.8633	690	2.8388	0.0245	0.00060
830	2.9191	810	2.9085	0.0106	0.00011
1700	3.2304	200	2.3010	0.9294	0.86382
760	2.8808	1190	3.0755	-0.1947	0.03792
1280	3.1072	710	2.8513	0.2560	0.06551

Step 1 through 3 are the same as for example 3A.

Step 4. Add the differences together (column 6) and divide by 2n, where n = the total number of pairs of duplicates (for this example n = 30) to get 0.09006.

Step 5. Take the square root of the result in step 4; this equals the pooled intralaboratory reproducibility standard deviation, which is 0.3001.

Step 6. To provide a higher range of values that cover what is likely to be observed, apply the coverage factor (k=2 for 95% coverage) to the intralaboratory reproducibility standard deviation to get the estimate of the expanded uncertainty of 0.600. (Note this is a log_{10} value).

Step 7. To calculate the uncertainty for any given result, the result is first converted to the log_{10} value and then the expanded uncertainty 0.600 is added and subtracted from the log result.

Step 8. To estimate the MU of a sample, convert the log value for the sample measurement back to base units (CFU) for the reported result. This is accomplished by taking the anti-log of the endpoints of the interval (anti-log of $x = 10^x$).

For example, for a sample result of 150 CFU: 150 in log_{10} = 2.1761 so the interval in log_{10} counts is 1.5761 to 2.7761; transforming to counts: $10^{1.5761}$ =37.7, and $10^{2.7761}$ = 597.2. Therefore, the uncertainty interval is 37 to 598 CFU.

DISCUSSION OF EXAMPLES 1-3:

The intervals produced in examples 2 and 3 can appear to be similar, but they are quite different and the calculations for estimating uncertainty are very different. In example 2 (recovery replicates), the main observation is the recovery on each trial and the components of uncertainty are reflected in the difference between different sets of replicates. A single SD is calculated as the estimate of combined standard uncertainty. In example 3 (intralaboratory reproducibility replicates) the components of uncertainty are reflected in the differences between replicates, so the difference between each pair of log counts is converted into a variance (squared difference) and these variances are pooled and then converted into an estimate of the combined standard uncertainty.

Though the procedure used for evaluating MU in Examples 1-3 are quite different, there are some similarities between the procedures. Whether the replicates are derived from control data replicates, as recovery replicates, or

as intralaboratory reproducibility replicates, they all include many of the same components of uncertainty listed in Table 5 below. Since the process in Example 1 only compares data over time, but not against another component as do the second and third examples, it is expected that the MU would be larger using the formulas described in Example 1. Since all three procedures are valid, and no one process is favored over the other, it is up to the laboratory to determine whether the MU estimate obtained from any of these processes is reasonable. The laboratory may consider combining these estimates taking into consideration the "double-counting". The laboratory should consider whether it meets the needs of its customers.

USING MPN DATA

According to Robert Blodgett who stated in the FDA BAM Appendix 2: Most Probable Number from Serial Dilutions 2010 "The MPN is particularly useful for low concentrations of organisms (<100/g), especially in milk and water, and for those foods whose particulate matter may interfere with accurate colony counts. Only viable organisms are enumerated by the MPN determination. If, in the microbiologist's experience, the bacteria in the prepared sample in question can be found attached in chains that are not separated by the preparation and dilution, the MPN should be judged as an estimate of growth units (GUs) or colony-forming units (CFUs) instead of individual bacteria. The following assumptions are necessary to support the MPN method: 1) The bacteria are distributed randomly within the sample. 2) The bacteria are separate, not clustered together, and they do not repel each other. 3) Every tube (or plate, etc.) whose inoculum contains even one viable organism will produce detectable growth or change, and 4) The individual tubes of the sample are independent".

It is important to note that the 95 percent confidence intervals in the MPN table (Table 1) has the following meaning: **Before the tubes are inoculated, the chance is at least 95 percent that the confidence interval associated with the eventual result will enclose the actual concentration**. Based on this statement, these confidence intervals should thus not be taken as the actual uncertainties.

Table 5: Potential Components of Uncertainty in Microbiology Analyses

Data	**Laboratory Control Samples**	**Type of Replicate** **Recovery**	**Type of Replicate** **Intralaboratory Reproducibility**	**Type of Replicate** **Intralaboratory Reproducibility**	**Method Validation**
MU Component	Example 1	Example 2	Example 3 a)	Example 3 b)	Example 4
Distributional					
Distribution of microbes in decision unit or primary sample (e.g. the total batch from which the laboratory sample was taken)	NA	NA	X Outside scope of this guide.	X Outside scope of this guide.	NA
Dispersion of microbes in laboratory sample	NA	NA	X	X	NA
Counting error	X	X	X	X	X
Technical					
Dilutions	X	X	X	X	X
Environment (Incubation chamber)	X	X	X	X	X
Equipment (Pipettors, balances)	X	X	X	X	X
Analyst	X	X	X	X	X

Data	Laboratory Control Samples	Type of Replicate Recovery	Type of Replicate Intralaboratory Reproducibility	Type of Replicate Intralaboratory Reproducibility	Method Validation
Different Days	X	X	X	X	X
Media Batches	X	X	X	X	X
Long Term (eg over a year)	X				
Standard uncertainty of spike concentration Standard uncertainty of a property value for a Certified Reference Material	X				
Other laboratory steps					
Matrix					
Laboratory Sample Matrix (inhibition or promotion or neutral)	X Caveat – use sample matrix		X	X Caveat – use sample matrix	X Caveat – use sample matrix
Preparation of Laboratory Control Sample (spikes into a clean matrix) A clean laboratory material that closely mimics the matrix of the material under test	X				
Trueness	NA		NA	NA	NA
Recovery (Bias)	X		X		

EXAMPLE 4 – USE OF METHOD VALIDATION DATA

This example is a way by which the estimate of reproducibility from an interlaboratory method validation study can be used as an estimate of measurement uncertainty. If a laboratory can demonstrate competence with the method, and if the uncertainty estimate is suitable for the laboratory's use, then no additional data need to be

generated (other than to estimate repeatability and bias, which should be done anyway). The procedure comes from ISO 21748: Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation.

The ISO document requires that the interlaboratory study to validate the method was conducted by competent laboratories using appropriate materials, and in accord with ISO 5725-2: Accuracy (trueness and precision) of measurement methods and results – Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method. This standard sets basic minimal numbers for laboratories, materials, and replicates, and has specific requirements for statistical analysis. Laboratories using measurement methods that are validated according to AOAC Harmonized Collaborative Laboratory Validation (HCLV) may also use this procedure for evaluating uncertainty.

The justification for this approach is that testing conditions in a single laboratory vary less over time than do conditions in different laboratories at a single point in time. Therefore, the estimate of reproducibility from a validation study provides a realistic over-estimate of uncertainty. If this estimate is acceptable for a laboratory, then the reproducibility estimate from the study will suffice. If it is not suitable, then laboratories can generate their own estimates as they gain experience with the method. The procedure requires that laboratories demonstrate that their competence with the method is consistent with the laboratories in the validation study. This is accomplished with checks for bias and repeatability, as described below. Finally, this procedure assumes that the validation study included all parts of the measurement procedure, including (for example) sample preparation. If this last assumption is not satisfied, then separate components can be added to the reproducibility, as discussed below.

Suggested Protocol for Evaluating MU using Statistics from a Validation Study

1. Assure there was an appropriate validation study design and data analysis (including outlier removal, statistical calculations, analysis for concentration effects, etc.), and that the estimates of Repeatability ($\mathbf{S_r}$) and Reproducibility ($\mathbf{S_R}$) are suitable for use in the laboratory.
2. Take the estimate of Reproducibility as a provisional estimate of measurement uncertainty (u′): $\mathbf{u' = S_R}$.
3. Use the estimates of Reproducibility and Repeatability to calculate the between laboratory SD ($\mathbf{S_L}$), as follows: $\mathbf{S_L = \sqrt{(S_R^2 - S_r^2)}}$.
4. Estimate the laboratory bias ($\mathbf{B_L}$) from repeated measurements of reference materials, comparison with a reference laboratory, or from proficiency testing: $\mathbf{B_L}$ = (Laboratory Mean – Reference Value). This should be calculated with log-transformed counts.
5. Estimate laboratory repeatability ($\mathbf{S_i}$) from an internal study (see discussion above about replicates), which could have been done in the past. The study should be based on at least 10 replicates.

   Note 1: if the laboratory estimate of bias ($\mathbf{B_L}$) is less than $\mathbf{S_L}$ then step 6 may be skipped. If repeatability $\mathbf{S_i}$ is smaller than $\mathbf{S_r}$, then step 7 may be skipped (but consider step 7c).

6. Calculate the acceptable criterion for the laboratory bias, as follows: **Bias Limit** $= \mathbf{2 \times S_L}$

   If $|\mathbf{B_L}| <$ **Bias Limit** then laboratory bias is acceptable for use of this procedure. If $\mathbf{B_L}$ is larger than the limit then the procedure cannot be used. Check the reason for the bias and correct it if possible.

7. Calculate the acceptance criterion for repeatability: **Precision Limit** $= \mathbf{1.5 \times S_r}$

   a. If $\mathbf{S_i <}$ **Precision Limit,** then repeatability is acceptable for this procedure

   b. If $\mathbf{S_i >}$ **Precision Limit,** then the procedure may still be used, but the provisional uncertainty estimate must be expanded as follows: $\mathbf{u' = \sqrt{(S_L^2 + S_i^2)}}$.

   c. If $\mathbf{S_i}$ is much less than $\mathbf{S_r}$ than the laboratory may wish to lower the provisional estimate of uncertainty using the same calculation in step 7.b, above.

Note 2: ISO 21748 actually requires a statistical F test here, based on the number of results in the validation study and the number of results used by the laboratory to estimate repeatability. This requires statistical complexity that is beyond this simplified description. For the minimum number of results in an acceptable study, and at least 10 true replicates in the repeatability study, the Precision Limit in this step is the tightest criterion that would be calculated.

8. Add in any components of variability ($\mathbf{S_{a1}}$, $\mathbf{S_{a2}}$, etc.) that were not included in the validation experiment, such as subsampling or sample preparation. There might be more than one

   additional component. Add the additional component(s) to the provisional estimate (**u′**) to produce the final estimate of combined standard uncertainty (**u**).

   $$\mathbf{u = \sqrt{(u'^2 + S_{a1}^2 + S_{a2}^2)}}$$

9. Calculate expanded uncertainty (**U)** with 95% coverage and **k** = 2, as follows: **U = 2xu**

   Note 3: If the uncertainty estimate is a percentage, the actual uncertainty will need to be calculated for every sample, as appropriate for its level.

   Note 4: Endpoints of the uncertainty interval are calculated with $\log_{10}$ values and transformed back to CFU.

Example: Using AOAC method 990.12: Aerobic Plate Count data

This method was validated by a study that used 8 laboratories, 6 foods with different levels of contamination, 2 samples per food, and 2 replicates per sample. The data analysis was consistent with ISO 5725-2, and the validation study included all steps in the testing process, except the step to choose an exact sub-sample size (measured samples were provided). Following are the reported estimates of Repeatability and Reproducibility for three of the foods, given as percentages:

Food	Reproducibility $\mathbf{S_R}$	Repeatability $\mathbf{S_r}$
Shrimp	11.1 %	9.8 %
Vegetables	9.2 %	6.3 %
Flour	5.8 %	5.3 %

Calculations described above in Steps 2, 3, 6 and 7 are used to generate **u′**, $\mathbf{S_L}$, **Bias Limit** and **Precision Limit**

Food	Provisional Uncertainty (**u′**)	Between Lab SD ($\mathbf{S_L}$)	Bias Limit ($2\mathbf{S_L}$)	Precision Limit ($1.5\mathbf{S_r}$)
Shrimp	11.1 %	5.2 %	10.4 %	14.7 %
Vegetables	9.2 %	6.7 %	13.4 %	9.4 %
Flour	5.8 %	2.4 %	4.7%	7.9 %

To estimate bias (Step 4), assume that the laboratory does a comparison study with a reference laboratory and results for vegetables and shrimp are always within 10% ($\mathbf{B_L}$ < 10.0 %). The comparison with a flour sample shows results 5% apart ($\mathbf{B_L}$ < 5.0%), so bias is judged to be acceptable.

To estimate repeatability (Step 5) the laboratory generates estimates with a series of 10 replicates, and repeatability for all foods is 5% or less ($\mathbf{S_i}$ < 5.0 %). It is decided, therefore, that repeatability is acceptable, but lower estimates of provisional uncertainty can be calculated, as described in Step 7c (using the formula in Step 7b):

Food	Initial Provisional Uncertainty (**u′**)	Laboratory SD ($\mathbf{S_L}$)	Repeatability ($\mathbf{S_i}$)	Final Provisional Uncertainty (**u′**)

Shrimp	11.1 %	5.2 %	5.0%	7.2 %
Vegetables	9.2 %	6.7 %	5.0%	8.4 %
Flour	5.8 %	2.4 %	5.0 %	5.6 %

In consideration of additional components (Step 8), we assume that sample preparation (sub-sampling, weighing) has been estimated (or is suspected) to add an additional 3.0% to the uncertainty (based on expert opinion). This component is then added as described in Step 8, and the Final Uncertainty is expanded as described in Step 9.

Food	Final provisional uncertainty (**u′**)	Adjusted final uncertainty (**u**)	Expanded Uncertainty (**U**)
Shrimp	7.2 %	7.8%	15.6%
Vegetables	8.4 %	8.9%	17.8%
Flour	5.6 %	6.4%	12.8%

Calculate uncertainty intervals for samples at 150 CFU as an example of how this is accomplished. The steps for this are described in procedures 1 and 2 above.

Food	Expanded Uncertainty (**U**)	Log_{10} of 150 CFU	Uncertainty in log_{10}	Uncertainty interval in log_{10}	Final uncertainty interval in CFU
Shrimp	15.6%	2.1761	0.3395	1.8366 to 2.5156	68 to 328
Vegetables	17.8%	2.1761	0.3873	1.7888 to 2.5634	61 to 366
Flour	12.8%	2.1761	0.2785	1.8976 to 2.4546	78 to 285

GENERAL CONSIDERATIONS FOR ESTIMATES OF UNCERTAINTY

Trueness is defined in the International vocabulary of metrology – Basic and general concepts and associated terms (VIM) as "closeness of agreement between the average of an infinite number of replicate measured quantity values and a reference quantity value". Trueness is the systematic error in a result. Trueness is often determined as bias using certified reference materials (CRM) with certified reference values (CRV). For many microbiological methods, trueness cannot be determined. This is because the analyte is a living organism and may be comprised of many different species, strains, etc. The condition of the organism may not be able to be defined, e.g., a heat-treated sample will contain organisms in different states of damage.

For microbiological laboratories these CRMs may not be available, so, bias cannot be determined. In the examples in this guide, trueness is not included. Not being able to calculate bias makes the microbiology methods similar to operationally defined methods. The measurement uncertainty applies to the reportable value as long as the method is followed exactly.

The Eurachem Guide: Terminology in Analytical Measurement – Introduction to VIM 3 states "such 'operationally defined' measurands are still fit for the purpose of comparing results and making decisions provided that the agreed measurement procedures are strictly followed."

If the laboratory must consider the uncertainty of the bias, the Eurachem Guide: Quantifying Uncertainty in Analytical Measurement (QUAM) provides instructions on how to do so.

Sampling is considered to be a large component of uncertainty in quantitative test results of any type. However, it is not considered by any of the procedures in this document. This is because for most measurement procedures, sampling is a separate process, and error due to sampling is usually not considered to be part of the uncertainty in the laboratory measurement. Some procedures include sampling as a part or measurement uncertainty, but most procedures do not. Sampling error is best studied separately, for purposes of control of that important component.

Estimates of uncertainty should always be confirmed if possible and the primary means is through professional or experienced judgment which should always be recorded as to acceptability: The calculated uncertainty intervals should agree with expert experience for typical samples and should meet the needs of the customer. Proficiency testing can also provide useful confirmation of uncertainty estimates. Any result graded as "unacceptable" should be investigated to see if it is also outside calculated uncertainty interval, and then investigate appropriately.

Estimates of measurement uncertainty that are obtained consistent with procedures described in ISO 19036:2019 - Microbiology of the food chain - Estimation of measurement uncertainty for quantitative determinations are also acceptable. Examples 1, 3b and 4 in this document are consistent with ISO 19036. Example 3a is also consistent with ISO 19036 if appropriate organisms are used for the spike.

DOCUMENT REVISION HISTORY

Date	Description
07/23/19	➢ Complete reorganization and editorial revisions

Contributors: A2LA wishes to acknowledge the efforts of Dawn Mettler, Dan Tholen, Jane Weitzel, Brad Stawick, Shawn Kassner and Roger Brauninger in the development and editorial input of this document.