

GAP ENVIROMICROBIAL SERVICES

A DISCUSSION OF THE UNCERTAINTY ASSOCIATED WITH THE BACTERIAL COLONY COUNTS OBTAINED BY MEMBRANE FILTRATION

(Revision 8 – March 10, 2004)

Prepared by: G. Horsnell
Microbiologist at GAP
EnviroMicrobial Services

A DISCUSSION OF THE UNCERTAINTY ASSOCIATED WITH THE BACTERIAL COLONY COUNTS OBTAINED BY MEMBRANE FILTRATION

TABLE OF CONENTS

SECTION	PAGE
SECTION 1 - INTRODUCTION.....	1
SECTION 2 - A REVIEW OF BASIC STATISTICAL CONCEPTS.....	4
SECTION 3 - WITHIN-ANALYST VARIATION.....	5
Single Analyst Tests a Single Sample Repeatedly.....	5
Confidence Limits about the Mean.....	8
Small Data Sets.....	9
Additional Information about Repeated Tests on a Single Sample.....	10
Calculating Standard Deviation (SD) and Uncertainty for Duplicate Tests on a Series of Samples.....	11
Data Transformation.....	13
SECTION 4 – AMONG-ANALYST VARIATION AND COMBINED UNCERTAINTY.....	19
SECTION 5 – COMBINED UNCERTAINTY AND EXPANDED UNCERTAINTY.....	22
Background.....	22
Method Components (Factors) that Contribute to Variation.....	23
An Examination of MF Method Components.....	23
Two Ways of Collecting Data for Determining Combined Relative Standard Uncertainty (ru_c).....	25
Method 1 (Testing Among all Analysts).....	26
Method 2 (Between-Analyst Duplicate Testing).....	28
Calculating Combined Relative Standard Uncertainty (ru_c).....	30
Frequency for Determining Combined Relative Standard Uncertainty (ru_c).....	31
Calculating Expanded Uncertainty (U).....	31
Use of t as the Coverage Factor.....	32
SECTION 6 – REPORTING UNCERTAINTY.....	33
Calculations for Reporting.....	33
Confidence Limits Revisited.....	33
Policy for Reporting Uncertainty.....	34
APPENDIX 1 - WORKED EXAMPLES.....	35
Example 1 Using Method 1(Among-Analyst Testing) for Determining Uncertainty.....	35
Example 2 Using Method 2 (Between-Analyst Duplicates) for Determining Uncertainty.....	36
APPENDIX 2 - TESTING METHOD COMPONENTS FOR SIGNIFICANT SOURCES OF VARIATION.....	37
BIBLIOGRAPHY.....	44

A DISCUSSION OF THE UNCERTAINTY ASSOCIATED WITH THE BACTERIAL COLONY COUNTS OBTAINED BY MEMBRANE FILTRATION

SECTION 1

INTRODUCTION

The organization called the Canadian Association of Environmental and Analytical Laboratories (CAEAL) now requires laboratories to determine the uncertainty associated with analytical results from accredited methods. So, CAEAL requires procedures for determining combined and expanded uncertainty and a policy for reporting results with their expanded uncertainty.

However, in order to determine properly the amount of uncertainty associated with membrane filtration (MF) results, analysts should have an understanding of uncertainty and how the variation or uncertainty for the final MF bacterial results from a laboratory may be calculated.

The bacterial colony counts are seldom identical when an analyst uses MF to analyze a sample repeatedly. There is always some random variation in the colony counts among the replicate tests even though the best methods are used, all equipment is the same, the equipment is functioning properly and the analyst follows the procedure with great care.

The standard deviation (SD), the standard deviation squared (SD^2 or variance), the relative standard deviation (RSD) or the relative standard deviation squared (RSD^2 or relative variance) of the results each provides a different way of expressing the uncertainty (variation, fluctuation, spread, dispersion or scatter) associated with these repeated measurements under the same conditions.

However, the random variation in counts, which occurs when samples are analyzed under the same conditions, is not the only factor that contributes to the variation or uncertainty, which may be associated with any final result. When samples are analyzed in a laboratory, various factors or components of an MF method (e.g. sample holding time before analysis, batch of culture medium, filter lots, glassware, analysts who filter, analysts who count colonies, etc.) may change and contribute some additional variation to the results. So, the variation in counts associated with changes in these factors must be included when determining the uncertainty associated with any final result from a laboratory.

Furthermore, the variances that are associated with changes in various factors or components of a method are cumulative and must be combined before calculating the value of a statistic referred to as the combined standard uncertainty (u_c).

Combined standard uncertainty (u_c) may be calculated as the square root ($\sqrt{\quad}$) of the sum of all the variances or, since variance equals standard deviation squared (SD^2), it may be calculated as the square root of the sum of all of the squares of the standard deviations of the method components (factors), which independently contribute some uncertainty to the total (overall) uncertainty associated with the final result. Consequently, u_c is often called the root sum of squares and the following formulae can be used to calculate $u_c(y)$ or the combined standard uncertainty for the variable or measurand (y).

$$u_c(y) = \sqrt{[(\text{Variance}_{\text{FACTOR 1}}) + (\text{Variance}_{\text{FACTOR 2}}) + (\text{Variance}_{\text{FACTOR 3}}) \dots \text{etc.}]}$$

$$u_c(y) \text{ also} = \sqrt{[(SD^2_{\text{FACTOR 1}}) + (SD^2_{\text{FACTOR 2}}) + (SD^2_{\text{FACTOR 3}}) \dots \text{etc.}]}$$

It is also possible to use relative standard deviations (RSDs), where $RSD = SD/\text{mean}$, and determine combined uncertainty as combined relative standard uncertainty for the measurand (y), herein designated $ru_c(y)$ or simply ru_c . This may be calculated as the square root of the sum of all of the squares of the relative standard deviations. So, the formula for combined relative standard uncertainty is

$$ru_c(y) = \sqrt{[(RSD^2_{\text{FACTOR 1}}) + (RSD^2_{\text{FACTOR 2}}) + (RSD^2_{\text{FACTOR 3}}) \dots \text{etc.}]}$$

The standard deviation (SD) and the relative standard deviation (RSD) for each individual factor are referred to respectively as the standard uncertainty or the relative standard uncertainty for that factor.

However, the term u_c is called combined standard uncertainty because it combines the variation from all factors (method components or sources of variation) into a single, overall square root value equivalent to 1 standard deviation (SD) representative of the variation combined from all independent sources of variation.

The term ru_c is called combined relative standard uncertainty because it combines the variation from all factors (method components or sources of variation) into a single, overall square root value equivalent to 1 relative standard deviation (RSD) representative of the variation combined from all independent sources of variation.

We also know from basic statistical principles that, if a set of results has a normal frequency distribution, about 68% of results fall within ± 1 standard deviation of the mean, about 95% of results fall within ± 2 standard deviations of the mean and about 99% of results fall within ± 3 standard deviations of the mean where the values 1, 2 and 3 are referred to as expansion or coverage factors (k).

So, because u_c represents only 1 overall SD and ru_c represents only 1 overall RSD, they must be multiplied by an expansion or coverage factor (k), which is usually 2, to get 2 SDs, 2 RSDs or the expanded uncertainty (U) that will allow us to calculate the interval delimited by the mean ± 2 SDs or the range, which will include or cover about 95% of all values likely to occur around any estimate of the mean. Generally, this is the expanded uncertainty or range that we are trying to determine for any result produced by a laboratory.

A coverage factor (k) of 2 will suffice if the data set used to estimate u_c or ru_c includes a number (n) of observations or a number of degrees of freedom (usually $n - 1$) equal to or greater than 30. However, if small data sets are used for estimating u_c or ru_c and the number of degrees of freedom (df) is less than 30, laboratories will have to refer to t-tables for the Student's t distribution and replace the coverage factor (k = 2) with the value of t for the appropriate degrees of freedom at the 95% (or 0.05) level of confidence.

Laboratories should examine each of the various components (factors), which contribute to uncertainty for each method. However, when calculating combined standard uncertainty (u_c) or combined relative standard uncertainty (ru_c) and expanded uncertainty (U), the Canadian Association of Environmental and Analytical Laboratories (CAEAL) recommends including, in the estimate of combined standard uncertainty (u_c) or combined relative standard uncertainty (ru_c), only the variations (e.g. SDs^2 or $RSDs^2$) for those method components (factors) with standard deviations (SDs) that are 1/3 or more the size of the largest individual SD among independent method components.

During membrane filtration (MF), most of the variation in colony counts arises from differences in filtering procedures among analysts and differences in the ability of analysts to recognize target bacterial colonies (i.e. differences in colony counting among analysts). The variation (SDs) associated with

changes in other components will be small in comparison (i.e. SDs other components $< 1/3$ the SDs for filtering and colony counting among analysts) if quality control (QC) results show that equipment and materials are in control.

Sample holding time within the laboratory may be a significant source of variation for some tests. However, CAEAL will allow laboratories to exclude the variation associated with holding time if all samples are kept refrigerated and are analyzed within the holding time, which local regulations allow, between sampling and analysis.

In microbiology, we will use RSDs to determine the variation associated with various membrane filtration method components and use $RSDs^2$ to calculate combined relative standard uncertainty (ru_c) before determining expanded uncertainty (U). Furthermore, rather than examine each factor separately, there are at least two ways of collecting data that will already include the variation from many method components into the data, which can then be used for determining combined relative standard uncertainty (ru_c).

1. Occasionally, get all analysts to test the same sample. Calculate the RSD^2 for filtering among analysts and the RSD^2 for counting among analysts. Then, calculate ru_c as the square root of the sum of the two $RSDs^2$.
2. Collect between-analyst duplicate data over many samples where 2 analysts test (filter) each sample and the analysts each count the target colonies on their own filters. Then, calculate ru_c as the square root of the RSD^2 for the between-analyst duplicate counts.

If sufficient samples are tested and there are sufficient degrees of freedom (about 30), laboratories can multiply the combined relative standard uncertainty (ru_c), which has been determined from either one of these methods, by 2 to obtain the expanded uncertainty (U).

In summary, the variation or uncertainty associated with MF bacterial colony counts generally increases as the number of method components that might change during analysis increases. So, laboratories must include or combine the variations or uncertainties from all factors or method components, which are likely to contribute significant variation, into a single overall or combined estimate of uncertainty, which can be used to calculate the expanded uncertainty, which can then be used to calculate the limits within which the obtained laboratory result should fall with some generally agreed to level of confidence (e.g. 95% confidence).

The remainder of this document discusses issues related to the uncertainty associated with bacterial colony counts. It will discuss the variation in the results, which occurs when a single analyst repeatedly tests a sample (within-analyst variation). It will discuss data transformation. It will discuss repeatability, reproducibility and among-analyst variation. It will also discuss combined uncertainty and present procedures for collecting data and determining combined relative standard uncertainty (ru_c) as well as procedures for calculating and reporting expanded uncertainty (U). However, before we start determining the uncertainty associated with bacterial colony counts, let us review a few basic statistical concepts.

SECTION 2

A REVIEW OF BASIC STATISTICAL CONCEPTS

The mean (arithmetic average) is a measure of the central tendency of a set (sample) of repeated measurements.

The mean is defined as $(\sum x)/n$

Where, \sum = sum of

x = each result from the set (sample) of results

n = the number of results in the set (sample) of results

Variance is a measure of the spread, scatter, dispersion or variation in results about the mean of a set (sample) of results.

Variance is usually defined as $(\sum(x - \text{mean})^2)/n-1$

However, when using duplicate tests over a series of samples to examine variation, variance may be determined in two ways. The variance based on duplicates is the sum of the squares of the differences ($\sum D^2$) between duplicate values divided by two times the number (n) of duplicate pairs. It may also be calculated as the mean of all the variances between duplicate pairs.

Therefore, the variance for duplicates = $\sum D^2/2n$

The variance for duplicates is also = $\sum (\text{variance pair 1} + \text{variance pair 2} \dots\dots + \text{variance pair n})/n$

In either case, this will be the variance about the mean (average) of all duplicate values.

The standard deviation (SD) is the square root of the variance, which also means that variance is standard deviation squared.

As mentioned earlier, variation may be expressed in terms of standard deviation (SD) or as relative standard deviation (RSD). The relative standard deviation is SD/mean . In other words, it is the size of the standard deviation (SD) relative to the size of the mean. This is also called the coefficient of variation (CV). So, if a set of results has a mean of 26 with a standard deviation (SD) of 5.1, the RSD or CV will be $5.1/26 = 0.196$. However, the RSD is often expressed as a percentage of the mean. So, in the example above, the RSD% (CV%) is $5.1/26 \times 100 = 19.6\%$.

Furthermore, if you know the mean and the relative standard deviation (RSD) of a set of results, you can calculate the standard deviation (SD). It will be a percentage of the mean. In the case above, for example, the mean was 26 with an RSD% of 19.6%. So, the SD is 19.6% of the mean of 26 or 5.1 (rounded).

In addition, if a set of results has a normal frequency distribution, about 68% of results fall within ± 1 standard deviation of the mean, about 95% of results fall within ± 2 standard deviations of the mean and about 99% of results fall within ± 3 standard deviations of the mean.

SECTION 3

WITHIN-ANALYST VARIATION

Single Analyst Tests a Single Sample Repeatedly

Table 1 shows the results that were obtained when a single analyst ran 60 membrane filtration (MF) tests for total coliforms (TC) on a single water sample. Figure 1 shows a frequency histogram of the results.

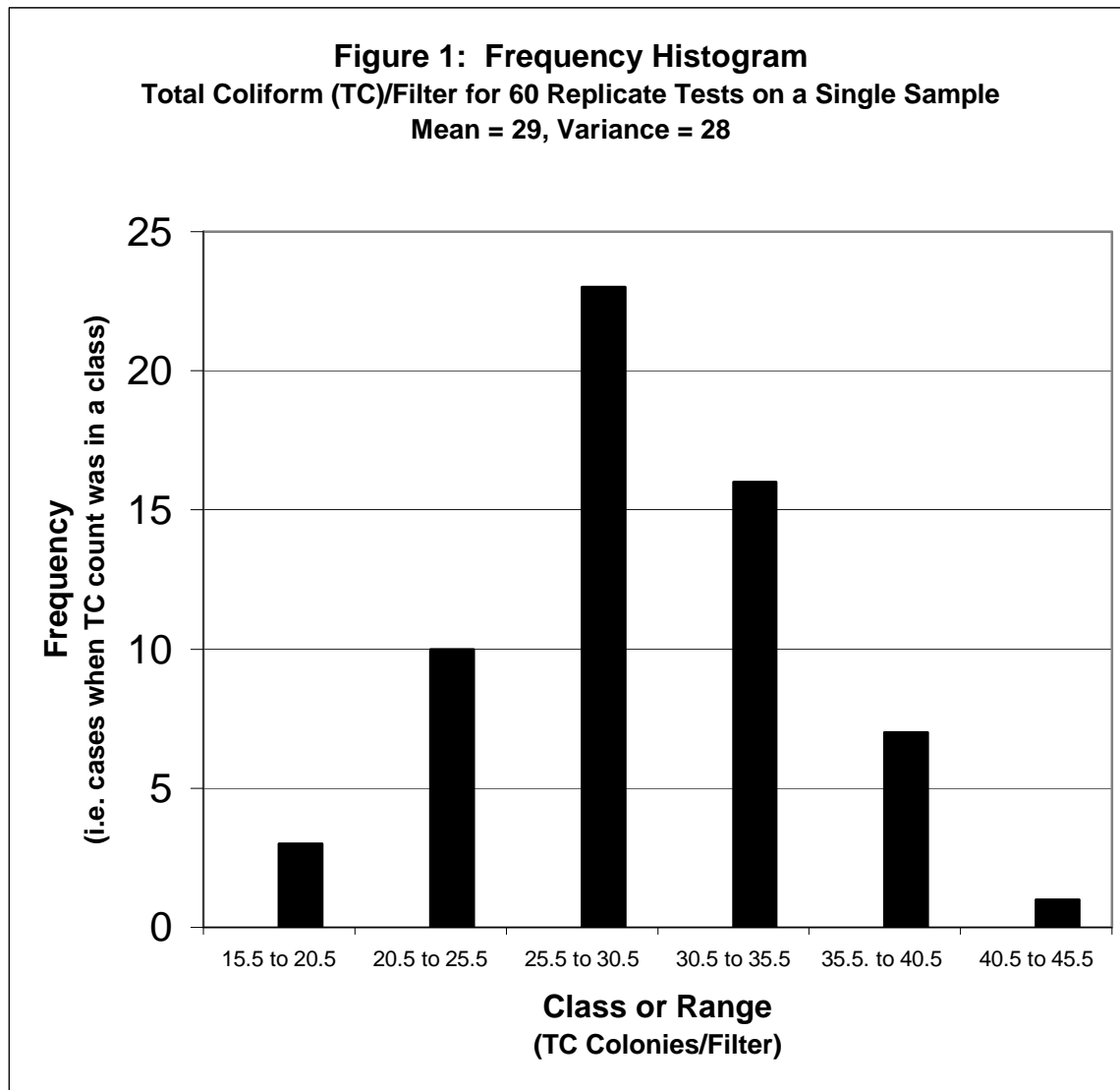
Table 1:

Colony Counts per Filter for 60 Replicate Membrane Filtration (MF) Tests on a Single Sample
 G. Horsnell, Ontario Ministry of Environment (MOE), Circa 1996
 TC = Total Coliform

Replicate	TC/Filter	Replicate	TC/Filter	Replicate	TC/Filter
1	24	21	30	41	29
2	27	22	31	42	24
3	36	23	25	43	27
4	27	24	34	44	18
5	43	25	38	45	32
6	21	26	26	46	32
7	28	27	31	47	23
8	30	28	34	48	40
9	36	29	31	49	34
10	26	30	31	50	29
11	29	31	25	51	22
12	29	32	26	52	32
13	23	33	29	53	36
14	40	34	28	54	35
15	29	35	26	55	30
16	26	36	28	56	36
17	33	37	34	57	20
18	26	38	33	58	25
19	29	39	31	59	19
20	27	40	21	60	32
		n	60		
		Mean	29		
		Variance	28		
		SD	5.3		
		RSD (CV)	$5.3/29 = 0.183$		
		RSD% (CV%)	18%		
		RSD ²	0.034		

NOTE: A test for normality did not detect a significant difference from a normal frequency distribution. In addition, notice that the mean and variance were almost equal.

Figure 1 shows a frequency histogram of the 60 replicate total coliform counts per filter from Table 1.



The above data was collected when a single analyst analyzed a single sample 60 times. So, the variation in colony counts per filter among replicates was based upon the following analytical conditions.

Factors Unchanged

- Sample
- Analyst who filtered
- Filtration funnel
- Batch of culture medium
- Incubator
- Incubation temperature per plate
- Incubation time per plate
- Analyst who counted colonies

Factors Which Changed

- Different time for each filtration
- Different pipette for each filtration
- Different membrane filter for each filtration

The histogram (Figure 1) of the 60 replicate results had a bell shaped appearance without any appreciable skew. In addition, the mean of 29 and variance of 28 were very similar. This suggested a normal frequency distribution with a large standard deviation (SD). But, reference texts tell us that a true normal distribution is comprised of continuous data, which can include fractions of values, not discrete data like colony counts, which are always whole numbers because you cannot have a fraction of a colony.

In fact, Niemella and others have shown that replicate, whole number, bacterial colony counts often follow a Poisson distribution, which by definition has a mean and variance that are equal. In these cases, the standard deviation of replicate colony counts is equal not only to the square root of the variance but also to the square root of the mean count because the mean and variance are equal.

So, assuming that the above results have a Poisson distribution, the standard deviation (SD) of the 60 replicate colony counts can be estimated as the square root of the mean count of 29 or $\sqrt{29}$ or 5.4. Furthermore, any single colony count is the mean of a set of results, which includes only one value. So, also assuming that replicate colony counts follow a Poisson distribution, the standard deviation of a single colony count (C) is \sqrt{C} and Niemella states that $1/C$ is the relative variance or relative standard deviation squared (RSD^2) of a colony count.

Nevertheless, the histogram of the above data did have a bell shape without any appreciable skew and a test for normality indicated that a normal distribution of results could not be rejected at the 95% (0.05) level of confidence. In other words, the frequency distribution of the 60 replicate counts was not significantly different from a normal distribution; so, for statistical analysis, the data may be treated as if it had a normal distribution.

Assuming a normal distribution, the standard deviation (SD) of the counts was 5.3, which is very close to the standard deviation of 5.4 calculated when we assumed a Poisson distribution. The relative standard deviation (RSD) of the counts was $5.3/29 = 0.183$. So, the RSD as a percentage of the mean (RSD%) is $5.3/29 \times 100 = 18.3\%$ and 2 times the RSD% or $2RSD\% = 2 \times 18.3\% = 36.6\%$.

Also, in a normal distribution with $n > 30$, about 95% of the replicate values should fall within the range of the mean $\pm 2RSD\%$ of the mean or within the range of the mean $\pm 2SD$, where the number 2 is the expansion or coverage factor used to calculate the expanded uncertainty or the range within which 95% of values will fall.

So, for the data set above, the expanded uncertainty at the 95% level of confidence or the range within which 95% of replicate values should fall is $29 \pm (36.6\% \text{ of } 29)$ or 29 ± 10.6 , between 18.4 and 39.6 TC/filter or between 19 and 40 TC/filer because you cannot have fractions of colonies.

In fact, 56/60 (93.3 %) or almost 95% of values did fall between the values 19 and 40 TC/filter once again supporting the contention that the distribution of the 60 replicate colony counts was approximately normally distributed.

Finally, analysts should understand that the estimate of uncertainty for the above data set was based solely on the specific conditions, which existed, at the time of analysis. It only included random variation plus the variation caused by compulsory changes in very few factors (e.g. pipettes, filters). However, if more factors could change (e.g. different analysts, different batches of the culture medium, different incubators, etc.), the amount of variation or uncertainty (e.g. variance or SD) in results would likely increase and this issue will be addressed later when we discuss the concepts of combined and expanded uncertainty.

Confidence Limits about the Mean

In the previous example, the mean TC count/filter of 29, determined from 60 replicates, was an estimate of the population mean or the mean, which would have been obtained if the entire population of all possible such replicates from the sample had been analyzed instead of just 60 replicates.

Since the mean of the sample of 60 results (sample mean) was only an estimate of the population mean, we are not sure exactly what the population mean actually is. However, from the available data, we can determine, with 95% confidence, the range within which the population mean is likely to fall.

This range or the confidence limits (CL) about the mean can be determined from the following formula.

$$CL = \text{mean} \pm t (SD/\sqrt{n})$$

Where, CL = 95% confidence limits about the mean

mean = the mean (average) of the sample of results (i.e. 29 in the example above)

t = the t value for the appropriate degrees of freedom at the 95% (0.05) level of confidence

SD = the standard deviation of the sample of results (i.e. 5.3 in the example above)

n = the number of results in the sample of results (i.e. 60 in the example above)

NOTE: The t value, from t-tables of the (Student's) t distribution, for n - 1 (i.e. 60 - 1 = 59) degrees of freedom (df) at the 95% (or 0.05) level of confidence is between 2.04 and 1.96 or about 2.

So, the following calculation will provide an estimate of the 95% confidence limits about the mean for the example above.

$$95\% \text{ CL} = 29 \pm 2(5.3/\sqrt{60}) = 29 \pm 2(5.3/7.746) = 29 \pm 1.37$$

This means that, based on the sample of 60 replicate counts, we can be about 95% confident that the population mean will fall somewhere within the range of 29 ± 1.37 , which is between 27.6 and 30.4 or between 28 and 31 colonies/filter because you cannot have fractions of bacterial colonies.

If a sample is analyzed repeatedly, it is useful to determine the confidence limits around the mean and to estimate, with 95% confidence, where the population mean is likely to fall. However, this is different from determining the mean ± 2 SD (or the mean ± 2 RSD% of the mean) or the range, which is likely to include 95% of all results.

For example, the confidence limits around the mean were from 27 to 31 TC/filter, which is the range quite likely to include the population mean. However, the range, which included about 95% of replicate results, was from 19 and 40 TC/filter.

This distinction is important because rather than analyzing samples repeatedly, most samples in microbiology are analyzed only once giving only a single value, which estimates the mean. So, without replicate results, we cannot determine confidence limits for the mean. However, we can use other data, which provides the variation for each of the many method components associated with use of the method within the laboratory, to determine the combined standard uncertainty (u_c) or the combined relative standard uncertainty (ru_c) and, then, use expanded uncertainty (U) to estimate the range within which 95% of colony counts would likely fall around any single estimate of the mean.

Small Data Sets

In a perfect normal distribution, 95% of values will fall between the mean ± 1.96 standard deviations, between the mean $\pm (1.96 \times \text{SD})$ or between the mean $\pm (1.96 \times \text{RSD}\%$ of the mean) where the expansion value of 1.96 is often rounded to 2.

In addition, when there are 30 or more observations, we can use an expansion value of 2 to calculate where 95% of values will fall around the mean.

So, for the above set of 60 replicate results that had a frequency distribution, which was approximately normal with a mean of 29, an SD of 5.3 and an RSD% of 18.3%, the range that would include 95% of all values was $29 \pm (2 \times 5.3)$ or $29 \pm (2 \times 18.3\% \text{ of } 29)$ or 29 ± 10.6 .

However, when the data set contains less than 30 values (i.e. $n < 30$) the distribution of the results will follow more closely a Student's t distribution (also known simply as a t distribution).

In these cases, about 95% of values will fall in the range the mean \pm (the t value for the appropriate degrees of freedom at 95% level of confidence \times SD). So, the value of t becomes the expansion or coverage factor used to calculate the range in which 95% of values will fall.

But, as the number of degrees of freedom ($df = n - 1$) for estimating SD decreases, t increases. For example, at the 95% (or 0.05) level of confidence, the t value for 30 degrees of freedom is 2.04 (approximately 2), the t value for 20 degrees of freedom is 2.09, the t value for 10 degrees of freedom is 2.23 and the t value for 5 degrees of freedom is 2.57.

NOTE: These t values can be found in standard t tables for the (Student's) t distribution in any elementary reference book on statistical analysis.

Now, let us assume that a single analyst tests a sample 6 times and gets the following results of 26, 29, 23, 30, 36 and 25 colonies per filter. This is a small set of < 30 results with a mean of 28, an SD of 4.6, an RSD of 0.164 and an RSD% of 16.4% where the degrees of freedom is $n - 1 = 6 - 1 = 5$. In this case, the t value at the 95% (or 0.05) level of confidence for 5 degrees of freedom is 2.57.

Because the data set is small, we must multiply the t value for the appropriate degrees of freedom (df) at 95% level of confidence \times SD in order to determine range around the mean, which will include 95% of values.

In other words, for this small data set, rather than using the mean $\pm 2 \times \text{SD}$, we must use the mean $\pm t \times \text{SD} = 28 \pm (2.57 \times 4.6) = 28 \pm 11.8$ to determine where 95% of values will fall around the mean.

Alternatively, to determine where 95% of values will fall around the mean, we can use the mean $\pm (t \times \text{RSD}\%$ of the mean) = $28 \pm (t \times 16.4\% \text{ of } 28) = 28 \pm (2.57 \times 4.6) = 28 \pm 11.8$.

Basically, 28 ± 11.8 is the expanded uncertainty at the 95% level of confidence for this small data set when a single analyst tested the same sample repeatedly. Furthermore, both the SD and the RSD% can be used to calculate the same estimate of expanded uncertainty at the 95% level of confidence.

Additional Information about Repeated Tests on a Single Sample

Table 2 summarizes the results obtained when a single analyst ran at least 60 MF tests repeatedly on 8 separate water samples analyzing each for a different analyte (e.g. total coliform, faecal coliform, etc.).

Table 2:

Results For 8 Trials of at Least 60 Replicate Membrane Filtration Tests Per Sample
G. Horsnell, Ontario Ministry of Environment (MOE), Circa 1996

Trial	Test	n	Mean Count/Filter	Standard Deviation	Variance	Range of Counts	Normality Test (95% conf)
1	TC	60	29.3	5.3	28	18 - 43	NSD
2	TC	60	3.97	2.0	4	0 - 9	NSD
3	FC	60	33.6	5.8	34	24 - 49	NSD
4	FC	60	21.2	6.6	44	6 - 37	NSD
5	FC	60	26.4	7.2	52	13 - 64	NSD
6	EC	64	28.5	6.9	48	16 - 44	NSD
7	EC	60	54.4	6.0	36	40 - 69	NSD
8	FS	60	9.1	2.7	7	4 - 16	NSD

n = Number of replicate tests per sample; TC = Total coliform; FC = Faecal coliform; EC = *E.coli*
FS = Faecal streptococcus; NSD = Not significantly different

The results show that, in many cases, the mean and variance of the colony counts per filter were similar, suggesting a Poisson distribution. However, analysis did not detect a significant difference from a normal distribution, at the 95% level of confidence, in any of the examples from Table 2. So, regardless of analyte, the replicate colony counts per filter can be analyzed as if they had a normal distribution.

Analysis also revealed that the relative standard deviation (RSD) of colony counts decreased as the counts per filter increased. For example, from Table 2, the RSD for TC was $2/3.97 = 0.5$ when the mean count was 3.97 per filter but the RSD for TC decreased to $5.3/29.3 = 0.18$ when the mean count was 29.3 per filter.

So, laboratories should determine and examine the variation or uncertainty associated with colony counts in the low (e.g. 0 - 19), middle (e.g. 20 - 80) and high (e.g. 81 - 150) range of target colonies per filter. These ranges were selected because it is generally agreed in microbiology that the acceptable counting range is 0 - 150 target colonies per 47 millimetre diameter membrane filter with an optimum counting range of 20 - 80 target colonies per filter. However, other ranges (e.g. 0 to 20% of total range, 21 to 50% of total range and 51 to 100% of total range) could be selected for uncertainty analysis.

In the cases above, a single analyst analyzed each sample repeatedly, so the variation or uncertainty was a measure of within-analyst variation or repeatability. An estimate of the variation associated with repeatability may also be obtained by calculating the variance for a series of different samples, which have been analyzed in duplicate by a single analyst (within-analyst duplicate data).

Calculating Standard Deviation (SD) and Uncertainty for Duplicate Tests on a Series of Samples

An estimate of uncertainty may be determined from the results obtained when a single analyst tests a series of different samples in duplicate. However, the duplicate data should be collected and analyzed within ranges and at least one filter of the pair should contain at least one bacterial colony.

Table 3 shows within-analyst total coliform (TC) duplicate data in the range of 20 - 80 colonies per filter. This is followed by an explanation of how to determine the variance SD, RSD, RSD^2 and uncertainty based on duplicate data.

Sample	TC/Filter Duplicate 1	TC/Filter Duplicate 2	Absolute Difference (D)	Difference Squared (D^2)	Variance
1	46	45	1	1	0.5
2	55	45	10	100	50
3	47	41	6	36	18
4	23	18	5	25	12.5
5	23	23	0	0	0
6	34	38	4	16	8
7	50	54	4	16	8
8	14	21	7	49	24.5
9	33	43	10	100	50
10	69	61	8	64	32
11	77	78	1	1	0.5
12	26	24	2	4	2
13	63	62	1	1	0.5
14	42	38	4	16	8
15	42	48	6	36	18
16	36	41	5	25	12.5
17	21	21	0	0	0
18	25	21	4	16	8
19	22	32	10	100	50
20	20	21	1	1	0.5
21	52	61	9	81	40.5
22	22	24	2	4	2
23	29	23	6	36	18
24	22	26	4	16	8
25	31	30	1	1	0.5
26	53	42	11	121	60.5
27	66	51	15	225	112.5
28	66	50	16	256	128
29	39	22	17	289	144.5
30	55	40	15	225	112.5
n = 30	Mean Count = 39		Mean D = 6.2	$\sum D^2 = 1861$	Mean Variance = 31

In Table 3, the number (n) of duplicate pairs was 30. So, 2n is 60. The mean of all duplicate values (counts) was 39.

It was mentioned earlier that the variance based upon duplicate counts over a series of samples could be determined in two ways. The same variance, SD, RSD and RSD² will be obtained either way.

In the first case, variance = $\sum D^2/2n$.

In the second case, variance = $\sum (\text{variance pair 1} + \text{variance pair 2} \dots\dots + \text{variance pair n})/n$.

Table 4 shows that both methods for analyzing duplicate data will give the same results for variation when we apply the methods to duplicate data from Table 3.

Table 4:			
Statistics for Duplicate Total Coliform (TC) Counts in Table 3			
Based on Variance for Duplicates = $\sum D^2/2n$		Based on Variance for Duplicates = $\sum (\text{var pair 1} + \text{var pair 2} \dots\dots + \text{var pair n})/n$.	
Statistic	Value	Statistic	Value
Number of data pairs (n)	30	Number of data pairs (n)	30
2n	60		
Mean count	39	Mean count	39
$\sum D^2$	1861		
Variance ($\sum D^2/2n$)	1861/60 = 31	Mean Variance	31
SD	$\sqrt{31} = 5.6$	SD	$\sqrt{31} = 5.6$
RSD (SD/mean count)	5.6/39 = 0.14	RSD (SD/mean count)	5.6/39 = 0.14
RSD ²	0.0196	RSD ²	0.0196

So, either method is acceptable for determining variation for calculating uncertainty based on duplicates.

In the example above, the RSD was 0.14. So, the RSD% would be $(5.6/39) \times 100 = 14\%$ and 2 times the RSD% or 2RSD% would be 28%. When this is applied to a count of 29 colonies per filter, we would estimate that about 95% of results should fall within the range of $29 \pm 28\%$ of 29 or between 29 ± 8.1 , between 20.9 and 37.1 TC/filter or between 21 and 38 TC/filter (rounded).

This estimate of uncertainty is similar although not identical to the range estimated from replicate tests on a single sample from Table 1 (i.e. from 19 to 40 TC/filter). However, we must remember that these are only estimates of uncertainty based upon available data using different methods. These different estimates of where 95% of values are likely to fall around any particular estimate of a mean count will tend to converge and become more similar as more samples are tested and more data is collected.

To stray off topic a little, a control limit for within-analyst precision per range may be set if you multiply the mean absolute difference for duplicate results per range by 3.267 (sometimes rounded to 3.27). In the case above (Table 3), the mean absolute difference was 6.2. So, the control limit for within-analyst precision in the range of 20 - 80 colonies per filter is $6.2 \times 3.267 = 20$. In future, if the absolute difference between duplicate counts in this range is less than 20, within-analyst precision remains in control.

Data Transformation

When water samples are collected and analyzed over a period of time from the same site, the results may show some very high bacterial concentrations, some very low concentrations or some outliers compared to the majority of results because pollution conditions may change with time. Similar effects may also occur when a sample is tested at different serial, decimal dilutions. They may also happen when a solid food sample is analyzed repeatedly because it may be more difficult to homogenize a food sample than a water sample. So, there may be substantial skew in the frequency distribution of these results.

In these cases, the bacterial concentrations may have a lognormal frequency distribution where the logarithms (base 10) of the bacterial concentrations follow more closely a normal distribution than the untransformed data.

So, many microbiologists suggest that bacterial colony counts should be transformed or converted to the logarithm (base 10) of the counts before performing statistical analyses.

But, as shown earlier, when an analyst tested the same water sample repeatedly at the same dilution, the untransformed replicate colony counts per filter were approximately normally distributed and, when this occurs, it is not necessary to transform the data. In addition, it is not usually necessary to transform the duplicate counts per filter from within a short range of colony counts per filter.

Nonetheless, some analysts may feel safer transforming colony count data even though the untransformed results may already have a frequency distribution, which is approximately normal. In these cases, it will not hurt to transform the data. If the data is already approximately normally distributed, the estimates of uncertainty will be similar regardless of whether or not the counts are transformed to their logarithm. This is shown in Tables 5 and 6 for water sample duplicate results within a short range.

Table 5 shows, untransformed and log transformed, duplicate data in the range of 20 - 80 colonies per filter.

Table 5:					
Untransformed and Log Transformed Data for Duplicate Total Coliform (TC) Colony Counts					
GAP laboratory Data in Range 20 – 80 TC Colonies per Filter					
TC/Filter Duplicate 1	TC/Filter Duplicate 2	D ²	Log TC/Filter Duplicate 1	Log TC/Filter Duplicate 2	D ²
46	45	1	1.662758	1.653213	0.000911
55	45	100	1.740363	1.653213	0.007585
47	41	36	1.672098	1.612784	0.003518
23	18	25	1.361728	1.255273	0.011333
23	23	0	1.361728	1.361728	0
34	38	16	1.531478	1.579784	0.002333
50	54	16	1.69897	1.732394	0.001117
14	21	49	1.146128	1.322219	0.031008
33	43	100	1.518514	1.633468	0.013215
69	61	64	1.838849	1.78553	0.002864
77	78	1	1.886491	1.892095	0.000314
26	24	4	1.414973	1.380211	0.001208
63	62	1	1.799341	1.792392	0.000483
42	38	16	1.623249	1.579784	0.001889
42	48	36	1.623249	1.681241	0.003363
36	41	25	1.556303	1.612784	0.00319
21	21	0	1.322219	1.322219	0
25	21	16	1.39794	1.322219	0.005734
22	32	100	1.342423	1.50515	0.02648
20	21	1	1.30103	1.322219	0.000449
52	61	81	1.716003	1.78533	0.004806
22	24	4	1.342423	1.380211	0.001428
29	23	36	1.462398	1.361728	0.010134
22	26	16	1.342423	1.414973	0.005264
31	30	1	1.491362	1.477121	0.000203
53	42	121	1.724276	1.623249	0.010206
66	51	225	1.819544	1.70757	0.012538
66	50	256	1.819544	1.69897	0.014538
39	22	289	1.591065	1.342423	0.061823
55	40	225	1.740363	1.60206	0.019128

NOTE: When collecting duplicate data, at least one of the duplicate filters must contain at least one colony. Using duplicate results with zero colonies on both filters will not provide data useful for determining variation and will lead to unreliable estimates of uncertainty.

Table 6 presents an analysis of uncertainty based on the duplicate data in Table 5.

Table 6:			
Comparison of Statistics and Uncertainty for Untransformed Versus Log Transformed Duplicate Total Coliform (TC) Colony Counts from Table 5			
Range 20 - 80 TC Colonies per Filter			
Statistics and Uncertainty Based on Untransformed Data		Statistics and Uncertainty Based on Log Transformed Data	
Statistic	Value	Statistic	Value
n	30	n	30
2n	60	2n	60
Mean Count (C)	39	Mean Log Count	1.554043
ΣD^2	1861	ΣD^2	0.256
SD (dups)	5.6	SD (dups)	0.065
RSD	$5.6/39 = 0.14$	RSD	$0.065/1.554 = 0.04$
RSD%	14%	RSD%	4%
2RSD%	28%	2RSD%	8%
Uncertainty Range	$C \pm 28\% C$	Uncertainty Range	$\text{Log } C \pm 8\% \text{ log } C$
95% uncertainty for a count of 39	$39 \pm 28\%$ or 28 to 50	95% uncertainty for a count of 39 Where, $\text{log } 39 = 1.591$	$1.591 \pm 8\%$ $= 1.591 \pm 0.127$ or 1.464 to 1.718 Antilog = 29 to 52

The analysis shows that, when the duplicate data is analyzed per range, it will not make much difference if the uncertainty is determined with or without converting the duplicate counts per filter to logarithms.

Tables 7 and 8 show that, if analysts wish to convert colony counts to their logarithm, they should not combine all duplicate data from the entire acceptable colony counting range of 0 - 150 colonies per filter. They should still analyze duplicate data within ranges. Otherwise, they may obtain unrealistic estimates of uncertainty.

In Tables 7 and 8, the mean variance is used to estimate uncertainty. However, as mentioned earlier, $\Sigma D^2/2n$ can also be used to calculate the variance, which can be used for determining the uncertainty from duplicates.

Table 7 shows untransformed and log transformed duplicate data, which covers the range from 0 - 150 colonies per filter.

Table 7:					
Untransformed and Log Transformed Data for Duplicate Total Coliform (TC) Colony Counts					
Data Lumped Together for the Entire Range 0 - 150 TC Colonies per Filter					
TC/Filter Duplicate 1	TC/Filter Duplicate 2	Variance	Log TC/Filter Duplicate 1	Log TC/Filter Duplicate 2	Variance
2	1	0.5	0.30103	0	0.045310
2	4	2	0.30103	0.60206	0.045310
1	2	0.5	0	0.30103	0.45310
4	3	0.5	0.60206	0.477121	0.007805
6	8	2	0.778151	0.90309	0.007805
8	5	4.5	0.90309	0.69897	0.020832
15	7	32	1.176091	0.845098	0.054778
5	3	2	0.69897	0.477121	0.024608
2	4	2	0.30103	0.60206	0.045310
12	16	8	1.079181	1.20412	0.007805
8	14	18	0.90309	1.146128	0.029534
6	4	2	0.778151	0.60206	0.015504
8	12	8	0.90309	1.079181	0.015504
1	2	0.5	0	0.30103	0.045310
9	2	24.5	0.954243	0.30103	0.213343
4	7	4.5	0.60206	0.845098	0.029534
7	4	4.5	0.845098	0.60206	0.029534
1	3	2	0	0.477121	0.113822
3	6	4.5	0.477121	0.778151	0.045310
1	5	8	0	0.69897	0.244280
36	39	4.5	1.556303	1.591065	0.000604
49	57	32	1.690196	1.755875	0.002157
74	61	84.5	1.869232	1.78533	0.003520
56	58	2	1.748188	1.763428	0.000116
100	101	0.5	2	2.004321	0.000009
123	110	84.5	2.089905	2.041393	0.001177
112	91	220.5	2.049218	1.959041	0.004066
103	108	12.5	2.012837	2.033424	0.000212
93	88	12.5	1.968483	1.944483	0.000288
96	93	4.5	1.982271	1.968483	0.00095

NOTE: When collecting duplicate data, at least one of the duplicate filters must contain at least one colony. Using duplicate results with zero colonies on both filters will not provide data useful for determining variation and will lead to unreliable estimates of uncertainty.

Table 8 shows the uncertainty, which will be obtained from the log-transformed data in Table 7.

Table 8:	
Statistics and Uncertainty Based on Log-Transformed Duplicate TC Colony Counts in Table 7	
When Data is Lumped Together for the Entire Range 0 - 150 Total Coliform (TC) Colonies per Filter	
Statistic	Value
n	30
Mean Log Count (Log C)	1.039308
Mean variance	0.036626
SD	0.19
RSD	$0.19/1.039308 = 0.18$
RSD%	18%
2RSD%	36%
Uncertainty Range	Log C \pm 36% log C
95% uncertainty for a count of 102 Where, $\log 102 = 2.009$	$2.009 \pm 36\% = 2.009 \pm 0.72$ or from 1.289 to 2.729 (as logs) Antilog 19 to 536

When duplicate data over the entire range from 0 - 150 colonies per filter was lumped together and log transformed, analysis indicated that the uncertainty surrounding a count of 102 would be from 19 to 536 colonies per filter.

However, if an analyst obtained 102 colonies on duplicate 1, it is highly unlikely that the analyst would get either 19 or 536 colonies on duplicate 2 unless the analyst made a serious blunder during filtration.

Because the duplicate data was heavily weighted to low counts (i.e. 0 - 19 colonies per filter) and because the data was lumped together rather than separated into ranges, the precision or uncertainty of counts in the high range was overestimated even though the duplicate counts were converted to their logarithm.

So, as mentioned earlier, the data should be analyzed within ranges.

Table 9 shows that, if the data in the 81 - 150 range from Table 7 is analyzed separately, the estimate of uncertainty will be more realistic regardless of whether the data is log-transformed.

Table 9:

Statistics and Uncertainty Based on Duplicate MF Colony Counts from Table 7
 When Data from the Range of 81 - 150 TC Colonies per Filter is Analyzed Separately

Untransformed		Log Transformed	
Statistic	Value	Statistic	Value
n	6	n	6
Mean count (C)	102	Mean log count	2.00449
Mean variance	55.8	Mean variance	0.00097
SD	7.5	SD	0.03
RSD	$7.5/102 = 0.07$	RSD	$0.03/2.00449 = 0.15$
RSD%	7%	RSD%	1.5%
t(0.05), df 5	2.45	t(0.05), df 5	2.45
2.45RSD%	$2.45 \times 7.5/102 = 18\%$	2.45RSD%	$2.45 \times 0.03/2.00449 = 3.67\%$
Uncertainty Range	$C \pm 18\% C$	Uncertainty Range	$\text{Log } C \pm 3.67\% \text{ log } C$
95% uncertainty for a count of 102	$102 \pm 18\%$ $= 102 \pm 18.4$ or from 84 to 120	95% uncertainty for a count of 102 Where, $\text{log } 102 = 2.009$	$2.009 \pm 3.67\%$ $= 1.935 \text{ to } 2.083$ Antilog 86 to 121

In this case, each of the 6 samples had 2 results giving $2 - 1 = 1$ degree of freedom per sample or per data pair and the entire data set of 6 duplicate results had $6 \times 1 = 6$ degrees of freedom. So, the t value at the 95% confidence level for 6 degrees of freedom, which is 2.45, must be used to determine ± 2.45 RSD% of the mean or the range, which will include 95% of values.

The analysis shows that, when the untransformed data for the range 81 - 150 is analyzed separately, the estimate of uncertainty for a count of 102 will range from 84 to 120 colonies per filter. Using a log transformation, the estimate of uncertainty for a count of 102 will range from 86 to 121 colonies per filter.

Now, the estimates of uncertainty are similar, more reflective of the variety of differences between duplicate results, which occurred in this range, and closer to the estimate of $102 \pm (2 \times \sqrt{102})$ or between 82 and 122 where 95% of values would fall if we assumed Poisson scatter around a count of 102. Once again, a log transformation is not necessary.

SECTION 4

AMONG-ANALYST VARIATION AND COMBINED UNCERTAINTY

The previous section examined the variation (uncertainty), which occurred when a single analyst tested a sample repeatedly and when the analytical conditions remained essentially unchanged. So, the variance or the standard deviation (SD) of the replicate results (counts) provided a measure of an analyst's ability to repeatedly obtain the same result on a single sample, the within-analyst variation or the uncertainty associated with repeatability. However, it did not include any additional variation or uncertainty, which could occur if method components, such as analysts, could change.

The variance or the SD of the results (counts) obtained, when different analysts test the same sample or test samples in duplicate between analysts, provides a measure of the ability of different analysts to reproduce the results obtained by another analyst, the among-analyst variation or the uncertainty associated with reproducibility.

The between or among-analyst variance will usually be greater than within-analyst variance. This can be seen, in Table 10, when 11 different analysts used membrane filtration to analyze the same sample in duplicate for total coliforms (TC).

Table 10:				
Results When 11 different Analysts Analyzed the Same Sample in Duplicate				
GAP Data in Range 20 - 80 Total Coliform (TC) Colonies per Filter				
Analyst	TC Count/Filter Duplicate 1	TC Count/Filter Duplicate 2	Within Analyst Mean	Within-Analyst Variance
1	68	74	71	18
2	81	70	75.5	60.5
3	65	69	67	8
4	73	64	68.5	40.5
5	68	71	69.5	4.5
6	73	67	70	18
7	57	61	59	8
8	59	57	58	2
9	62	63	62.5	0.5
10	72	74	73	2
11	77	65	71	72
	Among-Analyst Variance for set 1 = 56	Among-Analyst Variance for set 2 = 29	Mean Count Overall = 68 Range = 58 - 75.5	Within-Analyst Mean Variance = 21

NOTE: A test for normality did not reject a normal distribution of replicate results at the 95% level of confidence for either set of counts per filter among analysts.

Table 10 shows that, when 11 different analysts tested the same sample, they obtained results with different relative accuracy. In other words, the mean count per analyst differed among analysts compared (relative) to the overall mean count. In addition, the analysts obtained the results with different precision (variance or SD) between duplicates. Furthermore, the mean among-analyst variance of $(56 + 29)/2 = 43$ was a little more than twice the mean within-analyst variance of 21.

The above data shows that the variation (uncertainty) associated with results tends to increase in a laboratory setting where more than one analyst tests samples. This is not uncommon and it makes sense because another source of variation or uncertainty (i.e. performance among different analysts) has been added.

Moreover, it is important to understand this concept because laboratories may not know exactly who, among all analysts, could perform a particular part of the test or the entire test procedure on a particular sample. So, they must include among-analyst variation when determining the uncertainty of results from the laboratory.

NOTE: The uncertainty may increase again if different laboratories test split samples particularly if different methods are used to perform the analyses for a particular analyte (e.g. total coliforms).

The data in the following Table 11 shows the variation that can occur when different analysts count colonies on the same filter or plate.

Table 11:			
Total Coliform (TC) Colony Counts/Filter Among Analysts for Different Ranges			
(GAP data where different analysts counted TC colonies on the same filter each time)			
Count for Analyst	Range 0 - 19 TC/Filter	Range 20 - 80 TC/Filter	Range 81 - 150 TC/Filter
1	14	61	127
2	14	57	100
3	14	58	136
4	14	61	101
5	14	67	100
6	14	60	116
7	14	64	102
8	14	61	
9	13	67	
10		67	
Mean	14	62	112
Variance	0.1	14	221
SD	0.33	3.7	15
RSD	0.024	0.06	0.13

Notice the size of the RSD increases from the lowest to the highest range, which emphasizes again that the variation in membrane filtration results should be determined per range. This increase in variation occurs because analysts may miss or count some colonies twice as counts per filter increase, because some analysts may not recognize subtle positive target colony reactions on more crowded filters or because there is more overlap of colonies as counts per filter increase and some analysts may count overlapping colonies as only one colony while others may distinguish them as more than one colony.

In addition, as previously mentioned, in any laboratory that has more than one analyst, any one of the analysts may filter or count colonies. So, after examining the information in Tables 10 and 11, analysts should understand that, if they want to determine the uncertainty associated with testing within the entire laboratory when there is more than one analyst, they should use the variance, SD and RSD for counts between or among analysts (i.e. the variation associated with reproducibility) to obtain the RSDs², which

will be added to the formula for calculating combined uncertainty. Furthermore, they must combine the RSDs² for different method components, which contribute variation independently, as in the following example.

In Table 10, a single analyst counted all colonies to obtain the results. So, the variation among analysts for counts in the range of 20 - 80 colonies per filter represents only that for the filtration component (factor 1) of the analysis. The mean variance for filtering among analysts was 43 with a mean count of 68. The RSD was $(\sqrt{43})/68$ or $6.6/68 = 0.097$ with an RSD² of 0.0094.

However, in Table 11, the variation among analysts for colony counts, in the range 20 - 80 colonies per filter, represents only that for the colony-counting component (factor 2) of analysis. The RSD was 0.06 with an RSD² of 0.0036.

In these cases, the among-analyst filtering variation and the among-analyst counting variation were calculated independently from components of the MF method, which contribute uncertainty independently. So, the variations (RSDs²) must be combined to obtain the overall uncertainty for the overall analytical process of filtering and counting colonies.

In other words, the overall RSD = $\sqrt{[(RSD^2_{\text{FACTOR 1}}) + (RSD^2_{\text{FACTOR 2}})]} = \sqrt{0.0094 + 0.0036} = 0.11$, which is greater than the uncertainty (RSD) for either the filtration component (factor 1) or the counting component (factor 2) alone.

In fact, this overall RSD of 0.11 is an estimate of combined relative standard uncertainty (ru_c) for colony counts in the range of 20 - 80 colonies per filter. It is simply an estimate because it was calculated from very little data and it does not include the variation or uncertainty associated with changes in other factors or components of the method (e.g. use of different incubators, changes in batches of the culture medium, etc), which might add even more variation or uncertainty to the results.

Hopefully, however, the above information provides readers with an idea of how the variation or uncertainty in results increases when more than one analyst can test samples and a fundamental understanding of the concept of combined uncertainty.

In the situation above, the uncertainty for only two analytical components, filtration among analysts and colony counting among analysts, was included in the analysis of uncertainty. However, membrane filtration has many method components and the variation associated with additional components (factors) may need to be included when determining the combined relative standard uncertainty (ru_c) for the results of membrane filtration testing within a laboratory. The following Section 5 deals with the issue of combined uncertainty and expanded uncertainty in more detail.

SECTION 5

COMBINED UNCERTAINTY AND EXPANDED UNCERTAINTY

Background

The previous sections have shown that the results in microbiology are seldom identical when a single analyst uses membrane filtration (MF) to perform tests repeatedly on a single sample. There is always some random variation (fluctuation, dispersion, spread, scatter or uncertainty) in the bacterial colony counts among the replicate tests even though the best methods are used, all equipment is the same, the equipment is functioning properly and the analyst follows the procedure with great care.

The standard deviation (SD), the standard deviation squared (SD^2 or variance), the relative standard deviation (RSD, where $RSD = SD/\text{mean}$) or the relative standard deviation squared (RSD^2 or relative variance) of the results each provides a way of expressing the variation or uncertainty in these results.

However, the random variation in counts that occurs, when samples are analyzed under the same conditions, is not the only factor that contributes to the variation or uncertainty, associated with any final result. When samples are analyzed in a laboratory, various components of an MF method (e.g. analysts who filter and analysts who count colonies) may change and add more variation. In addition, the sample holding time before analysis, the batch of culture medium, filter lots, glassware, etc. may change and contribute additional variation to the results. So, the variation in counts associated with changes in these factors might have to be included when determining the uncertainty associated with any final result from a laboratory.

Furthermore, the variances for various components of a method are cumulative and must be combined before calculating the combined uncertainty.

As mentioned earlier, in microbiology, the RSDs of method components (factors) will be used to calculate the $RSDs^2$, which will be incorporated into the following formula for combined relative standard uncertainty (ru_c).

$$ru_c(y) = \sqrt{[(RSD^2_{\text{FACTOR 1}}) + (RSD^2_{\text{FACTOR 2}}) + (RSD^2_{\text{FACTOR 3}}) \dots \text{etc.}]}$$

However, when calculating combined relative standard uncertainty (ru_c) and expanded uncertainty (U), CAEAL recommends including, in the estimate of combined relative standard uncertainty (ru_c), only the variations (i.e. $RSDs^2$) for those method components (factors) with standard deviations (SDs) that are 1/3 or more the size of the largest individual SD among independent method components.

In addition, if the data for calculating ru_c has equal to or greater than 30 degrees of freedom (df), laboratories can multiply ru_c by an expansion or coverage factor (k) of 2 to calculate the expanded uncertainty (U). However, if the data has less than 30 degrees of freedom (df), laboratories should multiply ru_c by the appropriate t value (from t tables) to calculate expanded uncertainty (U).

Before calculating combined relative standard uncertainty (ru_c) and expanded uncertainty (U) laboratories should list the various method components, which might contribute significant variation, and examine these method components (factors).

Method Components (Factors) that Contribute to Variation

The greatest amount of uncertainty associated with bacterial colony counts may be caused by the variation associated with field effects, which may include the variation associated with sampling or the variation in bacterial levels resulting from different conditions during sample transport to the laboratory.

After the laboratory receives the sample, there are many other components (factors) associated with laboratory effects and analysis, which might contribute significantly to uncertainty.

However, laboratories are responsible for only that portion of uncertainty, contributed by the laboratory. Therefore, to get a good sense of the potential sources of variation, laboratories should list the components (factors), which contribute to the combined uncertainty associated with laboratory effects. For membrane filtration, the factors include but may not be limited to the following.

- Variation in the temperature of sample storage at the laboratory
- Variation in sample holding time within the laboratory before analysis
- Variation in different batches of a culture medium (e.g. mEndo agar) made from a particular lot
- Variation in different lots of the same culture medium (e.g. mEndo agar)
- Variation among brands of a particular culture medium (e.g. mEndo agar)
- Variation among different brands of filters
- Variation among different lots of the same brand of filter
- Variation in the accuracy of volume measuring devices (e.g. pipettes, graduated cylinders)
- Variation in the volume of liquid in dilution blanks
- Variation associated with serial dilution
- Variation in the filtering performance of any given analyst (within-analyst filtering variation)
- Variation in the filtering performance among analysts (among-analyst filtering variation)
- Variation in temperature within an incubator
- Variation in temperature between or among incubators supposedly at the same temperature
- Variation in reading (colony counting) for a single analyst (within-analyst reading variation)
- Variation in reading (colony counting) among analysts (among-analyst reading variation)
- Variation associated with colony confirmation

An Examination of MF Method Components

The following formula can be used to calculate combined relative standard uncertainty (ru_c).

$$ru_c = \sqrt{[(RSD^2_{\text{FACTOR 1}}) + (RSD^2_{\text{FACTOR 2}}) + (RSD^2_{\text{FACTOR 3}}) \dots \text{etc.}]}$$

In theory, we should determine the variation (RSD^2) for each individual factor or method component and add it to the formula.

However, it is important to note, that the uncertainty (i.e. RSD^2) for any given method component must not be added more than once to the formula. Otherwise, this could lead to an overestimate of the uncertainty associated with the result. Therefore, when determining the uncertainty for factors such as culture media, holding time, incubators, filter lots, etc., where filtration is used to collect data, the uncertainty associated with the filtering process (component) of the method must be removed from these uncertainty determinations. Otherwise, the uncertainty for filtration will be added repeatedly to the formula, which is not acceptable.

It is very labour intensive and time consuming to examine each method component separately, remove the uncertainty for filtration when necessary and determine whether or not the uncertainty for a component needs to be included in the formula. Furthermore, according to CAEAL, it is not necessary to include the RSD^2 for a factor in the formula if the SD for that factor is $< 1/3$ the largest SD among method components. So, let us see if the uncertainty for certain method components can be excluded or combined during the data collection process before we determine combined uncertainty.

To start, let us look at the various factors that are included in the filtering process and those, which are independent of filtering.

Factors Associated with Filtering

Variation associated with different filters
Variation associated with different batches of culture media
Variation associated with accuracy of measuring devices
Variation in dilution blank volumes
Variation associated with serial dilution
Variation associated with filtering among analysts

Factors Independent of Filtering

Variation associated with sample storage temperature in the lab
Variation associated with holding time before analysis
Variation associated with incubators
Variation associated with colony counting among analysts

MF results between or among analysts, from many samples over time, will include the variation (uncertainty) for different filters, different batches of culture medium, differences in glassware accuracy, differences in dilution blank volumes as well as the variation associated with serial dilution and different filtering techniques among analysts.

Therefore, if among-analyst filtration data is collected properly, the uncertainty for method components associated with filtering will already be included in the overall estimate of uncertainty (e.g. RSD^2) for filtration. So, it will not be necessary to examine separately each of the method components associated with filtering.

However, this data will not include the variation in counts associated with factors independent of filtering (e.g. variation in temperature during sample storage, variation in temperature between incubators, holding times, counting among analysts).

But, the variation in counts associated with some factors independent of filtering (e.g. fluctuations in temperature during sample storage, the variation in temperature within or between incubators) will be negligible compared to the uncertainty in counts associated with filtering among analysts (i.e. $< 1/3$ the SD for filtering) if quality control (QC) results show that the equipment is in control (i.e. temperatures are kept within allowable tolerances). Furthermore, the variation in counts associated with differences among batches of the culture medium will be negligible (i.e. $< 1/3$ the SD for filtering) if QC shows that culture media preparation is in control and that there are no significant differences in target colony recovery between or among batches of the medium. So, it will not be necessary to include separately the variation associated with these factors if they are in control.

The variation in target colony counts associated with differences in sample holding time within the laboratory before analysis may be significant particularly in the case of heterotrophic plate counts (HPC). But, CAEAL will allow laboratories to exclude within-laboratory holding time as a significant source of variation if the samples are kept refrigerated and are analyzed within the holding time, which local regulations allow, between sampling and analysis.

However, according to CAEAL, independent factors with SDs equal to or greater than 1/3 the largest SD, should be included in the calculation of combined relative standard uncertainty (ru_c). So, laboratories may wish to examine the variation (e.g. SD) associated with holding time within the laboratory and include the RSD^2 in the calculation for ru_c if the SD for holding time is significantly large (i.e. > 1/3 the largest SD).

The variation obtained, when different analysts count colonies on the same filter, is also independent of filtering. It may be a significant source of variation and it should be included in the calculation of combined relative standard uncertainty (ru_c) if the SD for counting among analysts is greater than 1/3 the largest SD among method components.

The SD for counting in the low range (0 - 19 colonies/filter), as shown in Table 11, is often less than 1/3 the largest SD. Nevertheless, laboratories may always include this RSD^2 associated with counting among analysts for the sake of consistency when calculating combined relative standard uncertainty (ru_c) per range.

Theoretically, laboratories should examine the uncertainty for each method component. However, if all equipment and materials are in control, the greatest sources of variation within the laboratory and the factors, which really need to be considered, are the variation in counts associated with differences in filtration technique among analysts and the variation associated with the ability of analysts to recognize target colonies (i.e. colony counting among analysts). The uncertainty for holding time is optional.

With this in mind, the workload can be reduced if laboratories collect data for filtering and counting among analysts in a way that already includes the variation from other relevant factors. The uncertainties for other method components will either be included in the data or they will be insignificant, compared to the variation associated with filtering and counting among analysts, if all equipment and materials are in control. The following subsection shows ways of collecting data in this manner.

Two Ways of Collecting Data for Determining Combined Relative Standard Uncertainty (ru_c)

Given the previous information, the following procedures provide two ways of collecting data and determining combined relative standard uncertainty (ru_c).

1. Occasionally, get all analysts to test the same sample. Calculate the RSD^2 for filtering among analysts and the RSD^2 for counting among analysts. Then, calculate ru_c as the square root of the sum of the two $RSDs^2$.
2. Collect between-analyst duplicate data over many samples where 2 analysts test (filter) each sample and the analysts each count the target colonies on their own filters. Then, calculate ru_c as the square root of the RSD^2 for the set of between-analyst duplicate counts. This will include the uncertainty of method components associated filtering and colony counting.

Collecting data is fairly easy if there are only two analysts in the laboratory. However, the situation becomes more complicated when there are 3 or more analysts.

How, do laboratories determine the RSDs² for filtering and colony counting among analysts or collect duplicate data when there are 3, 4, 5 or more analysts?

The following examples show how to collect the data. It should be collected for ranges (i.e. 0 - 19, 20 - 80 and 81 - 150 target colonies per filter for MF) and we will assume that a log (base 10) transformation of the colony counts is not necessary if the data is analyzed per range.

In addition, we will assume that there are five analysts in the laboratory.

NOTE: The data in the following tables may be organized into a spreadsheet program such as Excel for easy analysis.

Method 1 (Testing Among all Analysts)

On 5 or more separate occasions, get all analysts to test a single sample. Furthermore, get them to prepare dilutions of the samples so that the variation associated with dilution blanks, different measuring devices (e.g. pipettes) and serial dilution will be included. But, get one analyst to count the colonies on all filters. This will eliminate any variation associated with differences in counting among analysts and give the variation associated only with differences in the filtering process among analysts.

In addition, on 5 or more separate occasions, get all analysts to count target colonies on a single filter. This will provide the variation associated only with differences in target colony recognition and counting among analysts.

Repeat this procedure for each analyte (e.g. total coliform, faecal coliform, *E.coli*, HPC, etc.) and for colony counts in each range (i.e. 0 - 19, 20 - 80 and 81 - 150 target colonies per filter).

Tables 12 and 13 provide examples for total coliform (TC) in the range of 20 - 80 colonies per filter, show how to organize the data and determine the RSDs². This is followed by a calculation of combined relative standard uncertainty (ru_c).

Table 12:**Uncertainty for the Filtration Component Among Analysts**

Total Coliform (TC) in the Range 20 - 80 TC/Filter

(all analysts filtered the same sample each time but one analyst counted colonies on all filters)

Analyst	TC/Filter Sample 1	TC/Filter Sample 2	TC/Filter Sample 3	TC/Filter Sample 4	TC/Filter Sample 5	TC/Filter Sample 6	
1	38	46	50	50	68	74	
2	41	28	58	54	81	70	
3	31	26	42	50	65	69	
4	33	34	50	33	73	64	
5	23	30	58	52	68	71	
Variance	48	63	45	71	40	13	
		Overall Mean Count = 51					
		Mean Variance = 47					
		$SD = \sqrt{47} = 6.9$					
		$RSD = 6.9/51 = 0.135$					
		$RSD^2 = 0.018$					
		Degrees of freedom = 6 x 4 = 24					

Table 13:**Uncertainty for the Colony Counting Component Among Analysts**

Total Coliform (TC) in the Range 20 - 80 TC/Filter

(all analysts counted the colonies on the filter each time)

Count from Analyst	Sample 1 TC/Filter	Sample 2 TC/Filter	Sample 3 TC/Filter	Sample 4 TC/Filter	Sample 5 TC/Filter		
1	55	71	43	61	20		
2	57	68	46	57	25		
3	61	72	33	58	22		
4	57	75	56	61	21		
5	60	71	34	67	22		
Variance	6	6.3	89	15	3.5		
		Overall Mean Count = 51					
		Mean Variance = 24					
		$SD = \sqrt{24} = 4.9$					
		$RSD = 4.9/51 = 0.096$					
		$RSD^2 = 0.0092$					
		Degrees of freedom = 5 x 4 = 20					

NOTE: The variation in counts for total coliforms (TC) is often large because the TC colonies often show considerable variation in reaction and not all analysts recognize subtle positive reactions.

Use the following formula to calculate combined relative standard uncertainty (ru_c).

$$ru_c = \sqrt{[(RSD^2_{\text{FILTRATION AMONG ANALYSTS}}) + (RSD^2_{\text{COUNTING AMONG ANALYSTS}})]}$$

In this case, ru_c for the range 20 - 80 TC/Filter = $\sqrt{(0.018 + 0.0092)} = 0.165$. Remember, this is an RSD.

Also, repeat the above process per range for each analyte (i.e. total coliform, etc.).

Method 2 (Between-Analyst Duplicate Testing)

Method 2 uses between-analyst duplicate data to determine combined relative standard uncertainty (ru_c). However, collecting duplicate data becomes complicated when there are 3 or more analysts.

Nevertheless, the following procedure may be used.

Give each analyst an analyst number. In this example, there are 5 analysts numbered 1 to 5. Organize the analysts to perform duplicate tests between analysts on a regular basis but rotate the analyst pairs so that they perform duplicate testing in the following or similar manner.

Sample 1	(Analyst 1 and Analyst 2)
Sample 2	(Analyst 1 and Analyst 3)
Sample 3	(Analyst 1 and Analyst 4)
Sample 4	(Analyst 1 and Analyst 5)
Sample 5	(Analyst 2 and Analyst 3)
Sample 6	(Analyst 2 and Analyst 4)
Sample 7	(Analyst 2 and Analyst 5)
Sample 8	(Analyst 3 and Analyst 4)
Sample 9	(Analyst 3 and Analyst 5)
Sample 10	(Analyst 4 and Analyst 5)
Etc.	Etc.

When the rotation is complete start over.

Each time the analysts run duplicate tests get each analyst to run the filtrations on the sample and then get the analysts each to count the colonies on their own filters.

Use the procedure above for each analyte (i.e. total coliform, faecal coliform, *E.coli*, HPC, etc.)

Continue the process throughout the year. Over many samples, the mean of the variances of the pooled counts between analysts will include the variation associated with dilution blanks, different filters, differences in measuring devices (e.g. pipettes), serial dilution, differences in filtering techniques and differences in counting, which will occur on average among analysts.

Analyze the data when all analysts have been included and, there are at least 30 duplicate counts per range (i.e. in the ranges 0 - 19, 20 - 80 and 81 - 150 target colonies per filter). To get a more reliable estimate of uncertainty, laboratories may analyze the data after collecting between-analyst duplicate results for a year or so (assuming that this will provide more than 30 duplicates per range).

Table 14 shows how to organize between-analyst duplicate data and calculate the RSD^2 for a range. This is followed by a calculation of combined relative standard uncertainty (ru_c).

Table 14:					
Uncertainty Among Analysts					
Total Coliform (TC) in the Range of 20 - 80 TC/Filter					
(5 analysts tested samples in duplicate in rotation and counted target colonies on their own filters)					
Sample	Analyst Pair		TC/Filter Duplicate A	TC/Filter Duplicate B	Variance
	A	B			
1	1	2	50	60	50
2	1	3	41	28	84.5
3	1	4	25	34	40.5
4	1	5	36	44	32
5	2	3	40	31	40.5
6	2	4	66	74	32
7	2	5	53	35	162
8	3	4	35	42	24.5
9	3	5	64	51	84.5
10	4	5	49	57	32
11	Etc.	Etc.	Etc.	Etc.	Etc.
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
Etc.	Etc.	Etc.	Etc.	Etc.	Etc.
		Overall Mean Count = 46			
		Mean Variance = 58			
		$SD = \sqrt{58} = 7.6$			
		$RSD = 7.6/46 = 0.165$			
		$RSD^2 = 0.0272$			

Use the following formula to calculate combined relative standard uncertainty (ru_c), when Method 2 is used to collect data, because the uncertainties for various components associated with filtering and counting among analysts are already combined into the duplicate testing procedure.

$$ru_c = \sqrt{[RSD^2_{\text{BETWEEN ANALYST DUPLICATES}}]}$$

Therefore, for the example above, ru_c for the range 20 - 80 TC/Filter = $\sqrt{0.0272} = 0.165$. Remember, this is an RSD.

Also, repeat the above process per range for each analyte (i.e. total coliform, etc.).

NOTE: If there are more than 2 analysts, laboratories should rotate analyst pairs to gather between-analyst duplicate data when using Method 2 for determining combined standard uncertainty. Otherwise, they may not capture all the variation, which might occur among analysts in the laboratory.

NOTE: Some laboratories have analysts who only filter samples and other analysts who only count colonies on plates. In these cases, they can use Method 1 or, they can combine techniques from Method 1 (e.g. to obtain the RSD^2 for counting among analysts) and Method 2 (e.g. to obtain the RSD^2 for filtering between analysts if one analyst counts colonies on the duplicates).

Calculating Combined Relative Standard Uncertainty (ru_c)

Given previous information; assuming that QC data shows that all equipment and materials are in control and excluding the effect of sample holding time within the laboratory, the combined relative standard uncertainty (ru_c) for colony counts obtained by membrane filtration reduces to the following formula if Method 1 is used.

$$ru_c = \sqrt{[(RSD^2_{\text{AMONG ANALYST FILTRATION}}) + (RSD^2_{\text{AMONG ANALYST COUNTING}})]}$$

So, if the colony count is in the range of 20 - 80 colonies per filter, if the RSD^2 for among-analyst filtering in this range is 0.018 and if the RSD^2 for among-analyst counting in this range is 0.0092, the following calculation will give the combined relative standard uncertainty (ru_c).

$ru_c = \sqrt{(0.018 + 0.0092)} = 0.165$. This is an estimate of ru_c as an RSD for the range 20 - 80 TC/Filter.

If Method 2 (i.e. between-analyst duplicate testing) is used, combined relative standard uncertainty (ru_c) reduces to the following formula.

$$ru_c = \sqrt{[RSD^2_{\text{BETWEEN ANALYST DUPLICATES}}]}$$

In the example using method 2, $u_c = \sqrt{0.0272} = 0.165$. This is also an estimate of ru_c as an RSD for the range 20 - 80 TC/Filter.

NOTE: If laboratories feel that holding time is a significant source of variation, they may examine the variation associated with holding time (see Appendix 2). If laboratories wish to include this uncertainty, add the RSD^2 for holding time to the formula for combined relative standard uncertainty.

$$\text{So, } ru_c = \sqrt{[(RSD^2_{\text{AMONG ANALYST FILTRATION}}) + (RSD^2_{\text{AMONG ANALYST COUNTING}}) + (RSD^2_{\text{HOLDING TIME}})]}$$

Alternatively, $ru_c = \sqrt{[(RSD^2_{\text{BETWEEN ANALYST DUPLICATES)}} + (RSD^2_{\text{HOLDING TIME)}}]}$

Frequency for Determining Combined Relative Standard Uncertainty (ru_c)

Once the combined relative standard uncertainty (ru_c) has been determined adequately per range for any particular method (e.g. total coliform, etc.), these estimates of ru_c should not change significantly if all elements of the method remain basically the same and in control.

Even changes in staff should not significantly alter the original estimate of ru_c as long as new analysts are properly trained and are not allowed to tests samples until they can produce results, which are not significantly different from other trained analysts.

However, ru_c should be re-evaluated whenever there are major changes or modifications to a method or whenever a new method is introduced.

If laboratories are uncomfortable with the concept of using the original estimates of ru_c in perpetuity, they can re-evaluate ru_c at some regular interval of their choice (e.g. every 5 - 6 years).

Calculating Expanded Uncertainty (U)

When relative standard deviations squared ($RSDs^2$) are used for determining combined relative standard uncertainty (ru_c), it is obtained in terms of a single relative standard deviation (RSD).

So, ru_c must be multiplied by an expansion or coverage factor (k) to get the expanded uncertainty (U) and calculate the interval delimited by ± 2 SDs, which cover about 95% of values around any estimate of the mean.

A coverage factor $k = 2$ may normally be used, if the number (n) of values used for determining combined relative standard uncertainty is greater than 30.

However, it may be necessary to replace k with t (from t tables, for n-1 degrees of freedom at the 95% level of confidence) as the coverage factor if the number (n) of observations or degrees of freedom (df) used for determining uncertainty is less than 30.

In most cases, a coverage factor (k) of 2 will suffice. Therefore, multiply the combined relative standard uncertainty (ru_c) by 2 to get the expanded uncertainty (U).

In other words, the expanded uncertainty (U) = $ru_c \times 2$

In the example above, when Method 1 was used to calculate combined relative standard uncertainty (ru_c), ru_c was 0.165. So, the expanded uncertainty (U) = $0.165 \times 2 = 0.33$. This is the equivalent of 2 RSDs.

NOTE: Once the combined relative standard uncertainty (ru_c) and expanded uncertainty (U) have been determined per range per test or analyte, record these values in a table so that they can be used as a reference when there is need to report uncertainty.

Use of t as the Coverage Factor

Laboratories can use an expansion or coverage factor (k) of 2 to calculate expanded uncertainty (U) when the data set used for determining combined relative standard uncertainty (ru_c) has 30 or more degrees of freedom (df).

However, when uncertainty is estimated from small data sets of less than 30 values, about 95% of values will fall in the range of any estimate of the mean \pm (the t value for the appropriate degrees of freedom at 95% level of confidence \times SD). This is based on the fact that the results from small data sets (< 30 values) follow more closely a Student's t distribution than a normal distribution. These t values can be found in standard t tables for the (Student's) t distribution in any elementary reference book on statistical analysis.

As the number of degrees of freedom (df) used for estimating SD decreases, t increases. For example, at the 95% (or 0.05) level of confidence, the t value for 30 degrees of freedom is 2.04 (approximately 2), the t value for 20 degrees of freedom is 2.09 and the t value for 10 degrees of freedom is 2.23.

Also, remember that the combined relative standard uncertainty (ru_c) represents 1 RSD. So, if ru_c was determined from a data set with 10 degrees of freedom, you would have to multiply $ru_c \times 2.23$ instead of 2 to obtain the expanded uncertainty (U) for estimating the range within which about 95% of values will fall around any estimate of the mean.

However, as a rule of thumb, you can probably use a coverage factor of 2 instead of t if the number of degrees of freedom is above 20, which has a t value of 2.09 at the 95% level of confidence.

For example, in method 1 for testing among analysts, 5 analysts tested 6 separate samples to examine the uncertainty associated with filtration. This provided 5 counts per sample for a total of 30 counts. Each set of 5 counts had $n - 1 = 5 - 1 = 4$ degrees of freedom. So, the number of degrees of freedom for the entire data set of 30 counts was $(n_1 - 1) + (n_2 - 1) + \dots + (n_6 - 1)$ or $6 \times 4 = 24$ degrees of freedom.

The t value for 24 degrees of freedom at the 95% level of confidence is about 2.06. So, this t value could have been used as the coverage factor. However, the t value of 2.06 is so close to 2 that using 2 as the coverage factor would not appreciably affect the estimate of expanded uncertainty. For example, when ru_c was 0.165, the results of $2.06 \times 0.165 = 0.339$ and $2 \times 0.165 = 0.33$ are not appreciably different.

In addition, different laboratories may have a different number of analysts. So, if they want to use a coverage factor of 2 rather than t , when using method 1 for determining ru_c , they will have to determine how many samples they must analyze in order to obtain a number of degrees of freedom greater than 20.

For example, if a laboratory has only 4 analysts, it could get all analysts to test each of eight samples, which will give $8 \times 3 = 24$ degrees of freedom with a t value of 2.06 at the 95% level of confidence. In this case, a coverage factor of 2 would probably suffice for calculating the expanded uncertainty (U). Alternatively, the laboratory could use the t value as the coverage factor or increase the number of samples tested. Nevertheless, remember that, if you want to use a coverage factor $k = 2$ confidently for estimating expanded uncertainty (U) at the 95% level of confidence, it is best to ensure that the data set has at least 30 degrees of freedom.

For between analyst duplicates (method 2), we suggested 30 data pairs per range. Each pair has $2 - 1 = 1$ degree of freedom. So, 30 samples will provide $30 \times 1 = 30$ degrees of freedom, which is large enough so that a coverage factor of 2 rather than the t value can be used for calculating expanded uncertainty (U).

SECTION 6

REPORTING UNCERTAINTY

Calculations for Reporting

Let us assume that 10mL of a sample is analyzed by membrane filtration for total coliforms (TC) and the colony count on the filter is 50 TC colonies, which is in the range of 20 - 80 colonies per filter. Let us also assume that the combined relative standard uncertainty (ru_c) for this range is 0.165 in terms of a single relative standard deviation (RSD). So, the expanded uncertainty (U) = $k \times ru_c = 2 \times 0.165 = 0.33$, which is 2 RSDs.

However, remember, if we multiply an RSD by 100 we will obtain the RSD as a percentage of the mean (i.e. RSD%). So, ru_c as RSD% = $ru_c \times 100 = 0.165 \times 100 = 16.5\%$ of the mean. So, U as RSD% = $2 \times 16.5\% = 33\%$ of the mean. Notice that this is actually the same as multiplying the expanded uncertainty (U) x 100 or $0.33 \times 100 = 33\%$. Furthermore, a single colony count is an estimate of the mean.

So, the expanded uncertainty (U) associated with this count of 50 or the range within which 95% of values are likely to fall will be the count $\pm 33\%$ of the count, which, in this case, is $50 \pm 33\%$ of 50 or 50 ± 17 (rounded) or between 33 and 67 colonies per filter.

However, the final result is normally reported as TC colony forming units (CFU) per 100mL. So, in this case, we must multiply the result per filter by 10. Therefore, the final result to the client may be reported as $500 \pm 33\%$ of 500 TC/100mL. Alternatively, the final result may be determined and reported in the following manner.

$$CFU \pm U/100mL = (\text{Count/Filter} \pm U \text{ for the range as RSD\%}) \times \text{Dilution Factor}$$

In this case, 10mL of sample was tested with a dilution factor of $100mL/10mL = 10$. So, the final result may be reported as $(50 \pm 33\% \text{ of } 50) \times 10 = (50 \pm 17) \times 10 = 500 \pm 170$ TC/100mL.

NOTE: The expanded uncertainty (i.e. $2 \times ru_c$ or U as RSD%) per range (e.g. 20 - 80) should be applied to any count/filter within the range. So, in the case above, the expanded uncertainty for a count/filter of 25 is $\pm 33\%$ of 25. For a count/filter of 55, it is $\pm 33\%$ of 55; for a count/filter of 75, it is $\pm 33\%$ of 75; etc.

Confidence Limits Revisited

Earlier, we mentioned confidence limits around the mean and provided the following formula.

$$CL = \text{mean} \pm t (SD/\sqrt{n})$$

In addition, we stated that these limits were usually calculated using the mean and standard deviation of multiple values from replicate testing.

However, in the example above, we have only a single result of 50 TC/filter for 10mL, which is the mean of a set of results containing only one value (i.e. $n = 1$). We also know that the combined relative standard uncertainty (ru_c) for counts in the range of 20 - 80 TC/filter was 16.5% as RSD%, which means that the standard deviation (SD) for a count of 50 is 16.5% of 50 or 8.25. Moreover, we know that the expansion

or coverage factor $k = 2$ can be used to replace the t value (if there are about 30 degrees of freedom for determining ru_c).

So, let's see what happens when we plug these values into the formula for confidence limits assuming that the single count of 50 is the mean, $k = t = 2$, $SD = 8.25$ and $n = 1$.

$$CL = 50 \pm 2 (8.25/\sqrt{1}) = 50 \pm 2 (8.25/1) = 50 \pm 16.5 \text{ or } 50 \pm 17 \text{ (rounded per filter).}$$

This is the count per filter \pm the 95% confidence limits around this count when 10mL of sample was tested. In this case, the dilution factor was 100mL/10mL = 10. Therefore, the result per 100mL would be $(50 \pm 17) \times 10 = 500 \pm 170$ with 95% confidence.

Notice that the range for the confidence limits around this single count or single estimate of the mean is the same as the range, which includes about 95% of all values likely to occur around this single estimate of the mean, based upon previous calculations using combined relative standard uncertainty (ru_c) and expanded uncertainty (U). This occurs because we only have 1 value as an estimate of the mean and $\sqrt{n} = \sqrt{1} = 1$.

In other words, combined relative standard uncertainty (ru_c) and expanded uncertainty (U) allow us to combine variation from all sources and estimate the confidence limits or the uncertainty around a single value or count when a single result is used to estimate the mean.

Policy for Reporting Uncertainty

The final results for total coliform, faecal coliform and *E.coli* \pm their expanded uncertainty, when required, should be reported per 100mL. Heterotrophic plate counts \pm the expanded uncertainty, when required, should be reported per millilitre.

The expanded uncertainty and level of confidence (e.g. 95%) should be reported with the final bacterial result (count) under the following circumstances.

- When requested by the client/customer
- When required by government legislation, regulation, limit, objective or guideline
- When required to verify that a method is suitable (fit) for its intended purpose

In addition, when a report of uncertainty is required, it should be accompanied by a short explanation of how the uncertainty was determined. For example, the expanded uncertainty (U) was determined from appropriate duplicate and/or other data, which included the uncertainties for relevant components of the analytical method but excluded holding time.

APPENDIX 1

WORKED EXAMPLES

Example 1 Using Method 1 (Among-Analyst Testing) for Determining Uncertainty

Let us assume that a raw water sample was collected and sent to the laboratory for total coliform (TC) analysis by membrane filtration onto mEndo agar LES. Let us also assume that the sample was analyzed as soon as it arrived at the laboratory, that there was no significant difference in TC recovery among batches of mEndo agar, that the laboratory had only one incubator for TC and that the total coliform colony count was 50 per 10mL without confirmation.

In this case, there was no holding time, the laboratory had only 1 incubator, there was no significant difference among batches of mEndo agar and the count is in the range of 20-80 colonies per filter. So, we only need to include the uncertainty (RSD^2) for filtration among analysts and the uncertainty (RSD^2) for counting among analysts for the range 20 - 80 colonies per filter to calculate combined relative standard (ru_c) and expanded uncertainty (U).

Step 1

Refer to RSD^2 calculations in this range for filtering among analysts. This was 0.018 (from Table 12). Refer to RSD^2 calculations in this range for counting among analysts. This was 0.0092 (from Table 13).

Step 2

Plug the $RSDs^2$ into formula for calculating combined relative standard uncertainty (ru_c).

$$ru_c = \sqrt{[(RSD^2_{\text{AMONG ANALYST FILTRATION}}) + (RSD^2_{\text{AMONG ANALYST COUNTING}})]}$$

$$ru_c = \sqrt{(0.018 + 0.0092)} = 0.165$$

The combined uncertainty as percent relative standard deviations ($RSDs\%$) = $0.165 \times 100 = 16.5\%$.

Step 3

Calculate the expanded uncertainty (U).

$$U \text{ (as } RSD\%) = ru_c \times 2 \text{ (as } RSD\%) = 16.5\% \times 2 = 33\%.$$

Step 4

Calculate the final result per reporting volume.

$$CFU \pm U/100mL = (\text{Count/Filter} \pm U \text{ for the range as } RSD\%) \times \text{Dilution Factor}$$

$$\text{In this case, } (50 \pm 33\% \text{ of } 50) \times 10 = (50 \pm 17) \times 10 = 500 \pm 170 \text{ TC}/100mL$$

Therefore, the TC count per 100mL is 500 ± 170 at the 95% level of confidence.

Example 2 Using Method 2 (Between-Analyst Duplicates) for Determining Uncertainty

Let us assume that the same conditions exist as in example 1, that 10mL of sample was filtered and that the total coliform colony count was 50 per filter without confirmation.

Step 1

Refer to RSD^2 calculations in this range for between-analyst duplicate testing. This was 0.0272 (from Table 14).

Step 2

Remember, duplicate testing will already include the uncertainties for various factors associated with filtering and counting among analysts.

So, plug the RSD^2 into formula for calculating combined relative standard uncertainty (ru_c).

$$ru_c = \sqrt{[RSD^2_{\text{BETWEEN ANALYST DUPLICATES}}]}$$

$$ru_c = \sqrt{(0.0272)} = 0.165$$

The combined uncertainty as percent relative standard deviations ($RSDs\%$) = $0.165 \times 100 = 16.5\%$.

Step 3

Calculate the expanded uncertainty (U).

$$U \text{ (as RSD\%)} = ru_c \times 2 \text{ (as RSD\%)} = 16.5\% \times 2 = 33\%.$$

Step 4

Calculate the final result per reporting volume.

$$CFU \pm U/100mL = (\text{Count/Filter} \pm U \text{ for the range as RSD\%}) \times \text{Dilution Factor}$$

$$\text{In this case, } (50 \pm 33\% \text{ of } 50) \times 10 = (50 \pm 17) \times 10 = 500 \pm 170 \text{ TC/100mL}$$

Therefore, the TC count per 100mL is 500 ± 170 at the 95% level of confidence.

APPENDIX 2

TESTING METHOD COMPONENTS FOR SIGNIFICANT SOURCES OF VARIATION

It is unlikely that the variation associated with fluctuations in temperature within refrigerators or incubators, that variations in temperatures between incubators (supposedly at the same temperature) or that variations among different batches of media will contribute significantly to the uncertainty associated with colony counts when QC results show that these method components are in control.

Nevertheless, the following information may be useful if laboratories wish to examine more closely the variation in colony counts associated with different components of the membrane filtration method in order to determine whether to include or exclude these components as significant sources of variation.

The Variation Associated with Different Batches of the Culture Medium

Changes in the batch, lot or brand of a particular culture medium, which is used to enumerate target bacterial colonies, may or may not lead to substantial or significant variation in colony counts.

However, if laboratories want to determine whether or not different batches of the culture medium will contribute significantly to uncertainty, use the following process.

Collect duplicate data in the range of 20 - 80 target colonies per filter for comparing new and old batches of the culture medium. This should be available from QC data. Then, analyze the data to determine whether there is a significant difference in target colony recovery among batches of the medium.

Collect at least 10 duplicate pairs. Then, use a paired t-test to determine whether or not a significant difference in target colony recovery can be detected as in the following example.

NOTE: It is preferable if a single analyst performs the duplicate tests and counts the colonies on both filters for each batch.

Table A2 - 1:

**COMPARISON OF TOTAL COLIFORM (TC) COUNTS ON NEW AND OLD BATCHES OF
mENDO LES AGAR USING A PAIRED t-TEST**
(GAP laboratory data)

BATCH	TC COUNT/FILTER		DIFFERENCE (d)
	NEW BATCH	OLD BATCH	
1	61	59	2
2	68	79	-11
3	71	57	14
4	76	69	7
5	56	71	-15
6	43	66	-23
7	67	61	6
8	54	74	-20
9	66	65	1
10	71	68	3
11	65	73	-8
12	70	76	-6
13	61	60	1
14	57	60	-3
15	22	30	-8
16	55	71	-16
17	31	21	10
18	66	78	-12
19	56	78	-22
Mean (Average) Difference (MD)			-5.26
Standard Deviation of Differences (SDd)			11.06
Calculated Paired t			-2.07 (SDND, 0.05)

Use the following formula to calculate the paired t value.

$$\text{Paired } t = \text{MD} (\sqrt{n}/\text{SDd})$$

In the example above, the paired $t = -5.26(\sqrt{19}/11.06) = 5.26 (4.36/11.06) = -2.07$ (absolute value 2.07) and there are $n - 1 = 19 - 1 = 18$ degrees of freedom (df).

The critical t (from t-tables) is 2.1 for 18 degrees of freedom (df) at the 95% (0.05) level of confidence.

The calculated t value was less than the critical t value for a two-tailed test. So, a significant difference in total coliform recovery among different batches of mEndo agar was not detected at the 95% level of confidence (i.e. Significant Difference Not Detected or SDND, 0.05). Therefore, there is no evidence so far that different batches of mEndo agar LES will give significantly different TC counts at GAP.

If a significant difference is not detected, it will not be necessary to include the uncertainty for the culture medium when determining combined relative standard uncertainty (ru_c).

However, if a significant difference in target colony recovery is detected among batches of the culture medium, it may be necessary to include the uncertainty for the culture medium when estimating the combined relative standard uncertainty (ru_c).

Because filtration is used when testing culture media, the media batches will have to be tested in the following manner and the uncertainty for the filtration component will have to be removed. Otherwise, the uncertainty for filtering will be included more than once in the estimate of combined relative standard uncertainty (ru_c), which is not acceptable.

Over a period of time, collect data for at least 10 batches (preferably 30) in the following manner and place the results into a table similar to Table A2-2 below.

NOTE: A single analyst should perform the replicate testing and count the colonies on all filters for each batch.

Table A2 - 2:						
UNCERTAINTY FOR A CULTURE MEDIUM						
(e.g. for total coliforms on mEndo agar LES)						
Batch	Within Medium		D^2	Between Medium		D^2
	New Batch Rep 1 Cnt	New Batch Rep 2 Cnt		New Batch Rep 1 Cnt	Old Batch Rep 3 Cnt	
1	61	59	4	61	49	144
2	68	79	121	68	47	441
3	71	57	196	71	47	576
4	76	69	49	76	59	289
5	56	71	225	56	31	625
6	67	61	36	67	51	256
7	65	73	64	65	83	324
8	57	60	9	57	70	169
9	31	21	100	31	11	400
10	71	68	9	71	58	169
Etc.	Etc.	Etc.	Etc.	Etc.	Etc.	Etc.
n	10			10		
Mean Count	62			56		
Sum D^2	813			3393		
$SD = \sqrt{(\text{Sum } D^2)/2n}$	6.4			13		
RSD%	10%			23%		
SD at a count of 50	10% of 50 = 5			23% of 50 = 11.5		
Variance (SD^2) at 50	25			132.5		

If the SD between batches is larger than the SD within batches, remove the uncertainty for filtering by calculating the square root of the between batch minus the within batch variance (SD^2) at a count of 50.

In this case, the SD for the culture medium will be $\sqrt{(11.5)^2 - (5)^2} = \sqrt{132.25 - (25)} = \sqrt{107.25} = 10.4$.

If the SD for the culture medium is equal to or larger than 1/3 the largest SD among method components, calculate the RSD² for addition to the formula for determining combined relative standard uncertainty (ru_c).

In the example above, the SD of 10.4 is very large and it is likely to be the largest SD or larger than 1/3 the largest SD among method components. So, calculate the RSD and RSD².

The RSD is the SD/mean = 10.4/50 = 0.2. Therefore, the RSD² for the culture medium (mEndo LES agar) will be 0.04.

Save the RSD². It will be used later in the formula for determining combined relative standard uncertainty (ru_c).

The Variation Associated with Sample Holding Time Within the Laboratory Before Analysis

NOTE: It will not be necessary to collect duplicate data for holding time if all samples are analyzed the day they arrive at the laboratory.

However, if the laboratory holds samples at 4°C overnight before analysis, this may lead to substantial or significant differences in colony counts compared to results obtained when the samples are analyzed the day they arrive particularly in the case of heterotrophic plate counts (HPC)

If laboratories want to determine whether or not holding time will contribute significantly to uncertainty, use the following process.

Collect duplicate data for counts in the range of 20-80 target colonies per filter to compare target colony counts per filter on a set of samples, which have been analyzed as soon as they enter the laboratory, and target colony counts on the same samples after they have been held for 24 hours at 4°C.

Next, use the same paired t-test procedure, outlined in Table A2 - 1, to determine whether or not there is evidence of a significant difference in target colony recovery between holding times.

If a significant difference is not detected, it will not be necessary to include the uncertainty for holding time when determining combined relative standard uncertainty (ru_c).

However, if a significant difference in target colony recovery is detected between holding times, it may be necessary to include the uncertainty for holding time when estimating combined relative standard uncertainty (ru_c).

Use the same procedure, outlined in Table A2 - 2, to collect more data, eliminate the uncertainty associated with filtering and determine the uncertainty associated only with holding time.

If the SD for holding time is > 1/3 the largest SD, save the RSD². It will be used later in the formula for determining combined relative standard uncertainty (ru_c).

The Variation Associated with Different Incubators Supposedly at the Same Temperature

NOTE: It will not be necessary to collect duplicate data if the laboratory only uses one incubator for a test.

However, if the laboratory has more than one incubator for a particular test, this may or may not lead to substantial or significant variation in colony counts if the incubator temperatures are not identical.

If laboratories want to determine whether or not the use of 2 different incubators will contribute significantly to uncertainty, use the following process.

Collect duplicate data in the range of 20-80 target colonies per filter to compare target colony recovery between the 2 different incubators.

Next, use the same paired t-test procedure, outlined in Table A2 - 1, to determine whether or not there is evidence of a significant difference in target colony recovery between incubators.

NOTE: If the laboratory has more than 2 incubators for a particular test, a paired t-test cannot be used. An analysis of variance will be necessary to compare target recovery among incubators.

If a significant difference is not detected, it will not be necessary to include the uncertainty for incubators when determining combined relative standard uncertainty (u_c).

However, if a significant difference in target colony recovery is detected between incubators, it may be necessary to include the uncertainty for incubators when estimating combined standard uncertainty (u_c).

Use the same procedure, outlined in Table A2 - 2, to collect more data, eliminate the uncertainty associated with filtering and determine the uncertainty associated only with incubators.

If the SD for incubators is $> 1/3$ the largest SD, save the RSD^2 . It will be used later in the formula for determining combined relative standard uncertainty (u_c).

The Variation Associated with the Temperature of Sample Storage

NOTE: It will not be necessary to collect duplicate data if the laboratory only uses only one refrigerator for storing samples before analysis.

However, if the laboratory stores samples in more than one refrigerator, this may or may not lead to significant variation in target colony counts.

If laboratories want to determine whether or not the use of 2 different refrigerators will contribute significantly to uncertainty, use the following process.

Collect duplicate data in the range of 20-80 target colonies per filter to compare target colony recovery between the 2 different refrigerators.

Next, use the same paired t-test procedure, outlined in Table A2 - 1, to determine whether or not there is evidence of a significant difference in target colony recovery between refrigerators.

NOTE: If the laboratory has more than 2 refrigerators, a paired t-test cannot be used. An analysis of variance will be necessary to compare target recovery among refrigerators.

If a significant difference is not detected, it will not be necessary to include the uncertainty for refrigerators when determining combined relative standard uncertainty (u_c).

However, if a significant difference in target colony recovery is detected between refrigerators, it may be necessary to include the uncertainty for refrigerators when estimating combined relative standard uncertainty (u_c).

Use the same procedure, outlined in Table A2 - 2, to collect more data, eliminate the uncertainty associated with filtering and determine the uncertainty associated only with refrigerators.

If the SD for refrigerators is $> 1/3$ the largest SD, save the RSD^2 . It will be used later in the formula for determining combined relative standard uncertainty (ru_c).

The Variation in Colony Counts for a Single Analyst (Within-Analyst Reading Variation)

It will not likely be necessary to determine this variation (uncertainty) if more than one analyst counts colonies. Among-analyst counting variation will exceed within-analyst counting variation and be more significant.

The Variation in Colony Counting Among Analysts (Among-Analyst Reading Variation)

There will likely be variation in the target colony counts among analysts when they count colonies on the same filter. This variation may not be significant when the number of target colonies per filter is small (e.g. in the range of 0-19 colonies per filter). However, the variation may be large and significant when the number of target colonies per filter is large (e.g. greater than 20 colonies per filter).

In order to determine the variation in counting, get each analyst to count target colonies on a single filter. Do this at some regular frequency (e.g. monthly) for each method and for counts per filter in each range (e.g. 0-19, 20-80 and 81-150) of target colonies per filter.

Place the counts in a table similar to the following and calculate the variation in the following manner.

Table A2 - 3: Uncertainty for the Colony Counting Component Among Analysts Total Coliform (TC) in the Range 20 - 80 TC/Filter (all analysts counted the colonies on the filter each time)					
Count from Analyst	Sample 1 TC/Filter	Sample 2 TC/Filter	Sample 3 TC/Filter	Sample 4 TC/Filter	Sample 5 TC/Filter
1	55	71	43	61	20
2	57	68	46	57	25
3	61	72	33	58	22
4	57	75	56	61	21
5	60	71	34	67	22
Variance	6	6.3	89	15	3.5
			Overall Mean Count = 51		
			Mean Variance = 24		
			$SD = \sqrt{24} = 4.9$		
			$RSD = 4.9/51 = 0.096$		
			$RSD^2 = 0.0092$		

Calculate the overall mean target colony count. In this case, the overall mean count = 51.

Calculate the mean variance.

In the example above, the mean variance is $(6 + 6.3 + 63 + 89 + 3.5)/5 = 24$.

NOTE: If the data sets are uneven, containing different numbers of analysts, use the formula for pooled variance (S_p) to obtain the mean variance. This formula can be found in most texts on statistical analysis.

Calculate the SD. In this case, the $SD = \sqrt{24} = 4.9$.

If the SD is equal to or greater than 1/3 the largest SD for a method component, calculate the RSD^2 for addition to the formula for calculating combined relative standard uncertainty (ru_c).

In the example above, the RSD is the SD/mean, which is $4.9/51 = 0.096$.

The RSD^2 will be 0.0092.

Repeat the above process for counts in the remaining ranges (i.e. 0-19 and 81-150).

NOTE: Variation may be significant even for low counts on tests like total coliform and faecal coliform where there is a lot of variation in target colony size, shape, colour or reaction. Furthermore, significant variation in counting among analysts may occur on any test when the colony counts are greater than 20.

The Variation Associated with Colony Confirmation

Most laboratories do not routinely run confirmation tests on target colonies obtained by membrane filtration. So, it will not be necessary to estimate the uncertainty associated with confirmation. Otherwise, refer to Niemella.

The Variation for Other Method Components

The variation associated with filters from different lots, volume measuring devices (e.g. graduated cylinders, pipettes, etc.), serial dilution and filtering among analysts will be incorporated into the variation determined using between or among analyst data from many samples.

BIBLIOGRAPHY

1. Canadian Association for Environmental and Analytical Laboratories (CAEAL). 2002. CAEAL on the Estimation of Uncertainty of Measurement in Environmental Testing. Version 1.4.
2. Ellison, S.L.R. et. al. (Editors). 2002. Quantifying Uncertainty in Analytical Measurement. Eurachem/CITAC guide. Second Edition.
3. Miller, J.C. and Miller, J.N. 1988. Statistics for Analytical Chemistry. Second Edition. Ellis Horwood Ltd., England.
4. Mills, W.J. 2002. Measurement Uncertainty in Environmental Analysis: Microbiology and Toxicity. CAEAL workshop. Toronto, Ontario.
5. Natrella, M. 1963. Experimental Statistics. National Bureau of Standards. Handbook 91
6. Niemella, S.I. 2002. Uncertainty of Quantitative Determinations Derived by Cultivation of Microorganisms. Centre for Metrology and Accreditation, Finland. Publication J3/2002.