

**Draft**

**Proposed Procedures for Estimating  
the Critical Level and Quantitation Limit**

**Consensus Group  
Committee I on Detection**

**for**

**Proposal to USEPA for Replacement of  
40 CFR, Part 136  
Appendix B MDL Procedure**

Procedure Number:  
Revision Number: 0

Date Initiated: 4/21/2006  
Effective Date: DRAFT 4

Date Revised: 5/11/2006  
Pages Revised: See Highlights

Total Number of Pages: 25

---

## 0.0 SCOPE

To provide a procedure by which an individual laboratory may derive accurate estimates of routine method sensitivity for most analytical procedures for which they report results. Short term, long term and on-going estimates,  $L_C$  (Critical Value) and QL (Quantitation Limit) are derived for uncensored methods as well as censored 2D and 3D methods. The sensitivity estimates are then used as benchmarks for accurate reporting of analytical test results.

## 1.0 INTRODUCTION

- 1.1 The Method Detection Limit (MDL) procedure as specified in 40 CFR Part 136 Appendix B defines the detection level as that level where one can be 99% certain that if the test method results indicate the presence of the analyte being evaluated (1% false positive error rate). Unfortunately, the MDL does not satisfy this definition in actual practice, because it is a poor estimate which does not take into account long term variability (e.g., varying instrument conditions), blank contamination or recovery bias. The MDL in theory is equivalent to the Critical Level ( $L_C$ ). The Minimum Level (ML) is considered a quantitation limit and is sometimes defined as 3.18 times the MDL. This definition yields 10% RSD at the MDL under ideal theoretical conditions, but in practice ideal conditions are rare and the ML often provides a poor estimate of the  $L_Q$ . Neither the MDL or the ML take into consideration bias or false negative error rate. The following procedure rectifies the shortcomings of the MDL/ML and provides a practical means to produce accurate and realistic estimates for  $L_C$  and  $L_Q$ .
- 1.2 The reporting of data below the quantitation limit is not always necessary, therefore an estimate for  $L_C$  is only required when data are reported below the quantitation limit. The need for reporting down to  $L_C$  for specific targets should be defined in the Data Quality Objectives for the project before it is initiated. The short term estimates,  $L_c$  and QL as described in this procedure could also be used to demonstrate proficiency for the initial start up of a method. Long term estimates,  $L_c$  and QL as described in this procedure will automatically replace the short term estimates.
- 1.3 This procedure is potentially applicable to analytical measurements that produce at least ratio scale data. Detection limits are not meaningful for nominal, ordinal or interval measurement scale data. A ratio level measurement scale has all the properties of an interval level scale plus a meaningful zero point. For example, temperature readings in degrees F and degrees C are interval level data (20°C is not twice as "hot" as 10°C), therefore detection limits would not be determined for F or C thermometers. It should also be noted that although this procedure applies to test methods capable of producing ratio scale data, it may not be necessary to determine the critical value and detection limit of a test method; the need to do so is inherently a function of the end use of the data. For example, if the quantitative range of a test method has been established, it would not typically be necessary to determine the critical value and detection limit if the test method were being exclusively used to monitor changes in the measurement variable that all fall well within the quantitative range of the method.
- 1.4 The estimate for the Critical Level  $L_C$ , as defined by this procedure, provides confidence that a result greater than  $L_c$  is different from the method blank, however the level of uncertainty is unknown for a single measurement between  $L_c$  and QL. That is, numerical values are not quantitative in this region. The numerical value of a single measurement is only known with a known level of certainty at or above the QL.
- 1.5 The distribution of data in the region of detection is generally assumed to be normal (Gaussian), therefore that assumption is also made for this procedure. If the user suspects that their data is non-normally distributed (Log-Normal for example), we recommend several techniques to evaluate its distribution for normality. If data is log-normally distributed, it is possible to perform a log transformation of the data and then apply the normal distribution techniques to the transformed data. A discussion of dealing with data which is neither normally nor log-normally distributed is beyond the scope of this procedure.

We recommend using the following tools to evaluate data distribution when you have at least 20 data points. Summary statistics such as the coefficient of variation or coefficient of skewness may be used, to give an indication that your data may not be normally distributed. Graphical techniques such as probability plots and box plots may be used to evaluate whether your data may be normally or log-normally distributed. All of the aforementioned techniques must be used with censored data. Formal tests to evaluate the normality of your data include the Shapiro-Wilk test ( $n \leq 50$ ), the Shapiro-Fancia test ( $n > 50$ ) or D'Agostino's test ( $n > 50$ ).

Shapiro, S.S. and M.B. Wilk, 1965. An analysis of variance test for normality (complete samples). *Biometrika*, 52:591-611.

Shapiro, S.S. and R.S. Francia, 1972. An approximate analysis of variance test for normality. *Journal of the American Statistical Association*, 63:1343-1372.

D'Agostino, R.B. 1971. An omnibus test of normality for moderate and large size samples. *Biometrika*, 58:341-348.

## 2.0 SUMMARY OF PROCEDURE

- 2.1 The specific procedure followed depends on the type of analytical technique being employed. For methods that typically produce a numerical result for a method blank (Uncensored Method), the results of between 7 and 20 blanks are used to statistically calculate a Critical Value and Quantitation Limit. For methods that typically do not produce a method blank with a numerical value (Censored Method), an estimate of the Quantitation Limit is derived based on the instrument "noise" level present in a method blank. Seven spiked replicates are then analyzed and a Critical Value calculated. The QL estimates for both Censored and Uncensored methods are then tested using spiked blanks, and if necessary, adjusted. The qualitative capabilities of the analytical method determine which Censored Method procedure will be followed. For a 2D method, the Quantitation Limit is set based on the quantitated value of the seven replicates. For a 3D method, the Quantitation Limit is determined by both the quantitative and the qualitative results of the seven replicates.
- 2.2 The initial Critical Value and Quantitation Limit results are initially based on very small data sets. Once enough data have been generated, the short term estimates of these values must be replaced by long term estimates. If the laboratory already has enough data to generate long term estimates the derivation of initial short term estimate may be skipped entirely. It should be noted that the initial, long term and on-going estimates for  $L_C$  and QL are all valid estimates. However, as estimates which are more representative of the laboratory process are generated they will replace the previous estimate. For example, once a long term estimate is generated it will replace, but not invalidate, the initial estimate and on-going estimates will update, but not invalidate, long term estimates.
- 2.3 The Critical Value and Limit of Detection are verified on an ongoing basis by the monitoring of method blanks and the analysis of QL check samples (QLCs) for both censored and uncensored methods. If the Critical Value and Quantitation Limit can not be verified, their estimates are revised appropriately to provide a continuous improvement in detection estimates.

## 3.0 DEFINITIONS

### Uncensored Method

Analytical methods that nearly always (at least 85% of the time) produce numerical values for method blanks (e.g., spectroscopic tests such as ICP-OES) are referred to as Uncensored Methods.

### Censored Method

Analytical methods that frequently ( $> 15\%$  of the time) produce non-numerical results for blanks (e.g., chromatographic methods such as GC and LC methods) are referred to as Censored Methods. Censored

Methods are additionally categorized as “two dimensional” (2D) and “three dimensional” (3D) techniques. Censored methods sometimes yield a near constant blank signal or a flat numerical result.

### **2D Method**

A 2D technique qualitatively identifies an analyte based on both a time dimension (retention time) and a response dimension (FID, ECD, TIC, UV, IR ...). Dual column confirmation, wavelength ratios, or dual detector methods, while providing additional conformational information are still considered 2D techniques because they still measure only in the time and response dimensions.

### **3D Method**

Relative to a 2D qualitative technique, a 3D technique provides additional qualitative information to identify the analyte of interest. The additional dimension of information includes, mass spectral information, FTIR spectral data and pattern recognition for the identification of multi-component analytes (e.g. PCB Aroclors by GC/ECD).

### **Critical Value, $L_C$ and $L_c$**

The Critical Level,  $L_C$ , is defined as the smallest amount or concentration of analyte that can be distinguished from a blank or zero at a high level of confidence; it is the smallest value at which a detection can be observed reliably. Conceptually,  $L_C$  is the smallest concentration that protects against false positives (Type I error) at a 1% error rate. This means that for every 100 measurements of a true blank only one value should fall above the  $L_C$ . While  $L_C$  represents the true Critical Level,  $L_c$  represents an estimate of the true Critical Level. For methods capable of reporting uncensored results, users may request that results less than  $L_c$  be reported. Detections may be reported to the critical value  $L_c$  but must be flagged as estimates when less than the QL. Results less than the  $L_c$  will be reported as non-detect or as "< QL" (e.g., or “QL U,” where the “U” qualifier indicates a non-detect).

### **ISO**

International Standards Organization

### **IUPAC**

International Union of Pure and Applied Chemist

### **Limit of Quantitation, $L_Q$ , $\hat{L}_q$ and Quantitation Limit, QL**

Limit of Quantitation,  $L_Q$ , is defined as the lowest concentration that, in the context of some level of precision and bias, meets all method identification criteria and produces quantitatively reliable results for the end use of the data. The default level of precision and bias for  $L_Q$  as defined by ISO and IUPAC is 10% RSD with no bias. While  $L_Q$  represents the true Limit of Quantitation, QL represents an estimate of  $L_Q$  as defined by this procedure. In this procedure LQ is the smallest concentration that protects against false negatives and satisfies specified tolerances for precision and bias. The  $\hat{L}_q$  is the initial limit of quantitation estimate which is used to derive the final QL.

### **MDL**

Method Detection Limit as previously defined by the USEPA in 40 CFR, Part 136 Appendix B. The MDL is a single laboratory short term estimate of  $L_C$ , where  $\alpha = 0.01$  following the "t" distribution and is typically based on six degrees of freedom.

### **Method Blank**

An unspiked or non-fortified reagent water sample which proceeds through the entire method, including all preparatory and determinative steps.

### **QLC**

The Quantitation Limit Check sample is a method blank (e.g., reagent water) or “clean” sample that is spiked at the QL with the analyte(s) of interest and processed through the entire analytical procedure to verify that such a spike will produce a detection.

#### **QLMS/QLMSD**

The Quantitation Limit Matrix Spike and Quantitation Limit Matrix Spike Duplicates samples are a blank real world sample matrix (sample with no detectable analyte(s) of interest) that is spiked at the QL with the analyte(s) of interest and processed through the entire analytical procedure to verify that such a spike will produce a detection.

## **4.0 PROCEDURE**

4.1 The procedure followed depends on whether or not the analytical test generates method blanks with numerical values. Methods that typically generate numerical values for method blanks, such as spectroscopic test, will use the Uncensored Method (section 4.1.1). Methods that typically do not generate numerical values for method blanks, such as gas and liquid chromatography procedures, will use the Censored Method (section 4.1.2). Alternatively, the larger of the Uncensored Methods procedure (4.1.1) or the Censored Methods procedure (4.1.2) may be used even when the analytical test generates numerical results for method blanks, but this may result in larger  $L_C$  and QL estimates. Refer to Attachment 8.1 which depicts the procedure in a schematic flow diagram.

### 4.1.1 Uncensored Methods

4.1.1.1 This procedure is used if at least 85% of the method blank analytes are reported as numerical values. For initial demonstration of performance (e.g., for a new analytical method), collect results for method blanks generated during routine operation of the method. The method blank must go through all the preparation and analysis steps of the method. A minimum of 7 method blank results is required in order to calculate an initial estimate of the detection limit. Each method blank should be processed in a different preparation batch.

**NOTE:** If it is necessary to initiate analysis immediately, an initial estimate of the critical value and detection limit may be made by analyzing seven blanks in a single batch. However, this short-term determination may underestimate routine variability. Replace the short term estimates by those determined using method blanks in a minimum of seven different batches as soon as they are available.

A method is defined as a unique combination of preparative and determinative steps. For example, if  $L_C$  and QL are derived for method 6010 (ICP/AES) using prep method 3010 (acid reflux digestion), separate estimates  $L_C$  and QL would be required if prep method 3015 (microwave assisted digestion) is substituted for prep method 3010.

4.1.1.2 If a larger number of method blanks are available, then they should be used to derive  $L_C$  and QL as long as there is no reason to suspect that a change in method sensitivity occurred during the time period in which the method blank data were collected. The objective is to accurately estimate the “true” (population) standard deviation,  $\sigma$ . The larger the number of blanks used to calculate the sample standard deviation,  $s$ , the better  $s$  estimates  $\sigma$  (i.e., the smaller the factor  $K$ ). However, it is not necessary to use more than several hundred (most recent) data

points, since additional improvements in the accuracy of the estimate  $\sigma$  will be small relative to the data gathering and computational effort. If the laboratory has existing method blanks results available, for a current method, they may be used for the detection limit determination, bypassing section 4.1.1 and proceed directly to section 4.2 of this procedure.

NOTE: This procedure will only be appropriate if numerical values are reported for each replicate blank used in the estimate. No more than 15% percent of the data set may contain blanks with non-numerical results and a minimum of seven numerical replicate measurements must be used in the calculation. Note that it is acceptable (and expected) that some method blank results will be negative values.

4.1.1.3 Method blank values known to be spurious errors that occurred during analysis should be discarded at the time of analysis, or where appropriate, corrected. The data set should consist of method blanks for which contamination was sufficiently small for the reporting of valid results to the data users. It is acceptable to apply statistical outlier tests, for example the Grubbs test (Attachment 8.3) to identify and discard anomalous method blank values for large historic data sets. However, for a small data set (e.g.,  $n < 15$ ) or for method blank data that have been recently acquired, a result should not be rejected solely on the basis of a statistical outlier test – a physical rationale for rejecting the result must be documented. Documentation of a physical rationale may be impractical for a large historical data set (e.g.,  $n > 100$ ) acquired over a long period of time. Under these circumstances, data may be rejected solely on the basis of statistical outlier tests, but this should be done with caution. The excessive rejection of method data will result in calculated detection limits and critical values that are biased low. Under no circumstances should more than 15% of method blanks be rejected.

4.1.1.4 Calculate the sample standard deviation,  $s$ , of the set of  $n$  method blank measurements

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

$x_i$  = measured value from analysis of  $i^{\text{th}}$  method blank ( $i = 1, 2, \dots, n$ )

$$\bar{x} = \sum_{i=1}^n x_i / n = \text{mean of } n \text{ replicate blank measurements}$$

4.1.1.5 Calculate an initial estimate of the critical level and quantitation limit using the following equations:<sup>1, 2, 3</sup>

$$L_c = \bar{x} + s K_{\gamma, p, v}$$

$$\hat{L}_q = \bar{x} + 3 s K_{\gamma, p, v}$$

**Note:** If  $\bar{x}$  is negative than do not include it in the calculation of  $L_c$  and  $\hat{L}_q$ .

When calculating the critical value and quantitation limit estimates several digits should be retained to prevent rounding errors. The final result should be rounded to two significant figures using conventional rounding rules. Table 1 lists values for  $K_{\gamma,p,v}$  for  $p = 0.99$  and  $\gamma = 0.01$  for various values of the degrees of freedom  $v$ .

For a set of  $n$  replicates,  $v = n - 1$ . Ideally (especially for analysis of inorganic analytes by spectroscopic methods), the mean of the set of  $n$  blank replicates should be near zero or much smaller than  $s K_{\gamma,p,v}$ .<sup>4</sup> A large positive sample mean that is significantly different from zero may be indicative of excessive blank contamination. Method blanks that were rejected due to high levels of contamination should not be included in the data set. Otherwise, include all blanks that are representative of routine laboratory operations in the data set used to calculate  $L_c$ .

If the sample mean is significantly less than zero, then investigate the method (e.g., instrument calibration) to determine if the negative bias can be corrected. The magnitude of the negative bias should be minimized to the extent that is practical ( $\bar{x} < s K_{\gamma,p,v}$ ). However, if the mean is found to be greater than zero it must be used in the equations to calculate  $L_c$  and  $\hat{L}_q$ .

4.1.1.6 When calculating quantitation limit and critical value estimates several digits should be retained to prevent rounding errors. The final result should be rounded to two significant figures using conventional rounding rules. Table 1 lists values for  $K_{\gamma,p,v}$  for  $p = 0.99$  and  $\gamma = 0.01$  for various values of the degrees of freedom  $v$ . For a set of  $n$  replicates,  $v = n - 1$ . Ideally (especially for analysis of inorganic analytes by spectroscopic methods), the mean of the set of  $n$  blank replicates should be near zero or much smaller than  $s K_{\gamma,p,v}$ .<sup>4</sup> A large positive sample mean that is significantly different from zero may be indicative of excessive blank contamination, inaccurate interelement correction factors or improper initial calibration. Method blanks that were rejected due to high levels of contamination should not be included in the data set. Otherwise, include ALL blanks that are representative of routine laboratory operations in the data set used to calculate  $L_c$ . This means a blank should not be excluded from the dataset unless the entire batch of data is also rejected. Method blanks should be run among and as normal samples. For example; if you don't normally proceed your samples by a blank or wash you should not do so with the method blank.

4.1.1.7 If multiple instruments are to be used for the same test and it is desirable to report a single detection limit and critical value for the set of instruments, then a minimum of 7 method blanks must be analyzed for each instrument and the standard deviation calculated for each instrument. For a set of  $r$  instruments, the standard deviation for the  $i^{\text{th}}$  instrument,  $s_i$  (where  $i > 1, 2, \dots, m$ ), is calculated using  $n_i \geq 7$  replicates analyzed on that instrument. If all the instruments possess similar sensitivities, then a pooled standard deviation,  $s_p$ , may be calculated as follows:

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \dots + s_m^2 (n_m - 1)}{(n_1 + n_2 + \dots + n_m - m)}}$$

- 4.1.1.8 When the standard deviation for each instrument is determined using a set of only seven replicates, the standard deviations may be pooled when the largest standard deviation is no greater than two (for single analyte methods) to three (for multi analyte methods) times the smallest standard deviation. Note that if the same number of replicates,  $n$ , is used for each instrument, the pooled standard deviation is simply the square root of the mean variance for the  $m$  instruments:

$$s_p = \sqrt{\frac{s_1^2 + s_2^2 + \dots + s_m^2}{m}}$$

The pooled standard deviation ( $s_p$ ) is substituted for  $s$  and the “degrees of freedom”  $v = n_1 + n_2 + n_m - m$  for the determination of  $L_c$  and  $\hat{L}_q$ .

- 4.1.1.9 If multiple instruments are used but the instruments possess significantly different sensitivities (i.e., the standard deviations are significantly different from one another using the F test in section 4.2.5), then establish separate detection critical values and detection limits for each instrument.<sup>5</sup> Alternatively subsets of instrument with similar sensitivities may be pooled. As opposed to calculating a pooled  $L_c$  and  $\hat{L}_q$  for a set of similar instruments, the highest critical value and quantitation limit may be used for the set of instruments, if this meets the measurement quality objectives for sensitivity.

- 4.1.1.10 After the quantity  $\hat{L}_q$  is calculated, process a series of at least seven replicate spikes (through the entire preparatory and determinative procedure) as near this calculated value as practical. The initial spike concentration must be within a factor of 1 to 2 the calculated value. For methods in which multiple analytes are simultaneously analyzed at least 90% of all the target analytes must be within 1 to 2 times the initial calculated value and no analyte may be greater than 3 times the calculated value. The relative standard deviation (as measured in terms of both the mean concentration and spike concentration) and the mean recovered analyte concentration must satisfy the following criteria:

$$\begin{aligned} (s_y / \bar{y}) \times 100 &\leq 20\%, \\ (s_y / L_q) \times 100 &\leq 20\%, \\ 50\% &\leq (\bar{y} / L_q) \times 100 \leq 150\% \end{aligned}$$

$s_y$  = standard deviation of set of replicate spike measurements

$L_q$  = known spike concentration for each replicate

$\bar{y}$  = mean of the measured concentrations of the set of replicate spikes

If above tolerances for precision and bias are not met, increase the spike concentration and analyze a second set of at least seven replicates until these tolerances are met. If multiple spike concentrations are processed, the lowest spike concentration that satisfies the above tolerances for precision and bias is the estimated quantitation limit, QL. If the above criteria cannot be met at any concentration, then all results from the method for that analyte must be reported as non-quantitative.

#### 4.1.2 Censored Methods

4.1.2.1 This procedure applies to analytical methods that often do not produce numerical values for blanks; that is, for methods that would produce non-numerical results for more than 15% of the method blank analyses. Use spiked replicates to generate sufficient data for censored methods. If the lowest possible detection limits are required, the level of the spike is critical– it needs to be high enough for reliable qualitative identification, but no higher, since variance usually increases with increasing concentration. Censored methods are additionally categorized as “two dimensional” (2D) and “three dimensional” (3D) techniques as defined in Section 3.

NOTE: A method is defined as a unique combination of preparative and determinative steps. For example, if  $L_c$  and QL are derived for method 8270 (semi-volatile GC/MS) using prep method 3510 (separatory funnel liquid/liquid extraction), separate  $L_c$  and QL estimates would be required if prep method 3520 (continuous liquid/liquid extraction) is substituted for prep method 3510.

4.1.2.2 Estimate the smallest analyte concentration that will produce a detected result and fortify a method blank at this concentration. As the term is used here, “detected” result is defined as a numerical value from a measurable analyte signal that is clearly distinguishable from background “noise” (i.e., the signal from a blank) under the routine operating conditions of the method which meets all qualitative method-specified identification criteria <sup>6</sup>. Typically, for methods that are readily amenable to evaluations of this nature, this concentration should produce an apparent analyte signal that is three to five times the apparent noise level (e.g., via a qualitative visual examination of chromatograms for GC and LC methods). Specific qualitative or quantitative criteria for analyte identification are usually required for 3D methods (ranges within ion abundance ratios must fall for mass spectroscopy methods) and may be required, though to a less extent, for 2D methods (e.g., the detection of the analyte peak within a specified retention time windows for the primary and confirmatory columns). When specified by a test method, the measured result produced at the selected spiking concentration must be sufficiently high to meet all the method-required criteria required for analyte identification. The laboratory may use prior experience (e.g., prior analytical data) or consideration of the signal to noise to determine this estimate.

4.1.2.3 The estimate must now be tested and verified. Analyze at least a single spiked blank at the estimated lowest concentration of reliable qualitative identification

through the entire analytical procedure (including all preparatory and determinative steps). If the analyte is not “detected” repeat the test at twice the initial spike concentration.

4.1.2.3.1 If the analyte is “detected” and the laboratory needs to demonstrate the ability to detect at a smaller concentration then the test may be repeated at half the initial spike concentration used in section 4.1 2.3.

4.1.2.3.2 Note the smallest concentration,  $x$ , at which “detection” was achieved.<sup>7</sup> The laboratory may analyze several different spike levels at the same time in order to quickly determine the smallest concentration providing “detection.” The spiking concentration for each instrumental system should be determined in this manner.

4.1.2.3.3 If multiple instruments are to be used to perform the same test and the user desires to use the same detection limit for all instruments, then the test of the detection limit estimate must be performed on each instrument, and the largest value of  $x$  from all the instruments may be used as the estimate.

4.1.2.4 Process (the preparatory portion of the method) a minimum of 7 replicates, each spiked at  $x$  (section 4.1.2.3.2). It is preferable to process the spikes in different analytical batches (e.g., to reduce the probability of under-estimating method variability). Analyze (the determinative portion of the method) the replicates on each instrument. Existing data, such as the variability estimate ( $s$ ) from low-level spikes previously analyzed, may be included.

4.1.2.5 The set of replicate spikes must satisfy the following tolerances for precision and bias:

$$\begin{aligned}(s / \bar{x}) \times 100 &\leq 20\%, \\(s / L_q) \times 100 &\leq 20\%, \\50\% &\leq (\bar{x} / L_q) \times 100 \leq 150\%\end{aligned}$$

$s$  = standard deviation of set of replicate spike measurements

$L_q$  = spike concentration for each replicate

$\bar{x}$  = mean concentration of the set of replicate spikes

If above tolerances for precision and bias are not met, increase the spike concentration and analyze a second set of at least seven replicates until these tolerances are met. If multiple spike concentrations are processed, lowest spike concentration that satisfies the above tolerances for precision and bias is the estimated quantitation limit, QL. Results should be reported as non-quantitative if the above criteria cannot be met at any concentration.

4.1.2.6 Separate critical values and quantitation limits should be calculated for each instrument. However, if the sensitivities of the instruments are similar (i.e., for the

determinative portion of the method), then a pooled standard deviation may be calculated as discussed in sections 4.1.1.8 through 4.1.1.10 and a single critical value and detection limit may be established for the set of instruments.

4.1.2.7 Estimate the critical level from the equation:

$$L_c = s K_{\gamma, p, v}$$

Verify that all the replicate spikes at the selected concentration ( $x$ ) produce measured concentrations greater than the calculated critical value and any method-specified ID criteria are satisfied and the tolerances for precision and bias in Section 4.1.2.5 are met.

4.1.2.7.1 For censored 2D methods, if one or more of the  $n \geq 7$  replicates (spiked at concentration  $x$ ) fail to produce a measured value greater than  $L_c$ , (e.g. or an instrumental response) the spike concentration should be increased until measured values greater than  $L_c$  can be consistently obtained. (A measured value greater than  $L_c$  is defined as a detection for 2D techniques.)

4.1.2.7.2 For censored 3D methods, if one or more of the  $n \geq 7$  replicates fails to produce any instrumental response or a result that satisfies all method-specified identification criteria, increase the spiking concentration until this occurs. It is not necessary to increase the spiking concentration to obtain a measured value greater than  $L_c$  unless the qualitative information provided by the method appears to be inadequate to confidently report the presence of the analyte. For example, a response greater than  $L_c$  would be required for a MS method when a target analyte did not possess any secondary ions or if only the Total Ion Chromatogram (TIC) response was being used, because in these examples the MS is being used as a 2D technique.

4.1.2.7.3 For both 2D and 3D analytical method, if a detection is not obtained for one or more replicates, analyze at least four replicates at the higher spike concentration and verify these replicates produce detections (e.g., measured results are greater than the calculated critical value  $L_c$ ) and satisfy the precision and bias criterion in Section 4.1.2.5. Establish this concentration as the quantitation limit, QL. The estimated quantitation limit, QL, is the lowest spiking concentration that consistently produces a detection (i.e., results greater than the calculated critical value or fulfillment of all method-specified analyte identification criteria). If the above criteria cannot be met at any concentration, then all results from the method for that analyte must be reported as non-quantitative.

4.1.2.8 Blank Check – There may be some analytes in certain methods that are frequently detected in method blanks (e.g., common laboratory contaminants, such as dichloromethane or acetone for VOCs) despite efforts to eliminate blank contamination. As appropriate for these analytes, establish a critical value and detection limit based on the method blanks as described in section 4.1.1. If numeric values are not available at least 85% of the time or an inadequate number

of blank results are available, use the procedure in section 4.1.2. Optionally, estimates may be calculated using both procedures (section 4.1.1 and section 4.1.2) and the greater of the two values utilized.

## 4.2 Development of Long-Term Estimates $L_c$ and QL

- 4.2.1 The short-term initial detection limit estimations performed according to sections 4.1.1 and 4.1.2 use small data sets ( $n < 20$ ), and they must be replaced with estimates using larger data sets once the data is available. Section 4.2 may be skipped if larger data sets ( $n \geq 20$ ) are initially available.
- 4.2.2 The quantitation limit check sample (QLC) spiked at the reported value of QL is analyzed with each batch of samples to establish the long-term estimate of  $L_c$  (for censored methods), to validate the QL and to demonstrate method precision and recovery on an on-going basis. The concentration of the quantitation limit check sample is initially determined from the initial demonstration of performance discussed in Section 4.1. During the period of time in which the additional data points are being collected (to obtain a total of at least 20 to 100 points needed for the long term estimate), each analyte in the QLC must be detected (e.g., no false negatives – no results below  $L_c$ ). The recovery of the QLC must be within 50% - 150% for all analytes, however recovery for up to 20% of the analytes in multi-analyte organic methods may fall in the range of 10% - 190%. The QLC is analyzed on a per batch basis as part of continuing demonstration of method performance.
- 4.2.3 The initial estimate is replaced once 20-100 data points (method blanks and/or QLCs) are available. If data is being pooled (sections 4.1.1.8 – 4.1.1.10) from multiple instruments of similar sensitivity following the same analytical procedure there must be a minimum of three data points per instrument. *Data should not be pooled if it is suspected that method sensitivity has significantly changed during the period of time over which data collection was performed. In particular, do not pool the data if the F test discussed below indicates method sensitivity has changed (section 4.2.5).* If at least 20 points are not available one year after the initial demonstration of sensitivity was performed, repeat the calculation of  $L_c$  using the available additional method blank data (or QLCs for censored methods) collected during this time. A total of 20 to 100 (ideally 100) of the most recent points must ultimately be available to determine long-term estimates of  $L_c$  and QL according to the equations and procedures specified in sections 4.1.1 and 4.1.2. Verify the new value of  $L_c$  using at least one spike at the QL as discussed in section 4.1.2.3.
- 4.2.4 The data set used must be representative of routine analysis. For example, if the detection limits are to be determined using method blank data then all method blanks must be included unless the laboratory noted an unusual contamination or analysis problem and rejected the samples in the associated batch. Limited outlier removal using standard procedures such as the Grubbs test may be performed. A method blank result greater than the initial estimate of  $L_c$  is not sufficient evidence per se to remove the blank result from the data set used for the long-term  $L_c$  estimate (a physical justification must also be documented).
- 4.2.5 Optionally, compare the variance of the new set of method blanks or QLSs with that of the original (the initial demonstration of performance) using a two-tailed F test

$$F = \frac{\left( s_H^2 \right)}{\left( s_L^2 \right)}$$

$s_H^2$  = Larger variance estimate

$s_L^2$  = Smaller variance estimate

If the calculated value of F is greater than the F-statistic value for the 95% level of confidence (or 99% level may be used when a large number of analytes is being simultaneously determined), then the critical value and detection limit must be updated. If the calculated value of F is less than the F-statistic value for the 95% or 99% level of confidence, then the data should be pooled to calculate a more reliable estimate of  $L_c$ . However, it should be noted that the F-test is not robust to departures from normality (e.g., the underlying assumption for the calculation of  $L_c$ ).<sup>10</sup> .

## 5.0 QUALITY ASSURANCE

Once long-term estimates have been established, monitor the validity of these estimates using on-going QC measures (method blanks and QLCs). Additionally, at a minimum of annually, verify that the proper procedures are being followed, estimates are still valid, and auditable documentation is available.

### 5.1 Ongoing verification of $L_c$ using method blanks

5.1.1 The Critical Level,  $L_c$ , as defined, should provide for 99% control of false positives at concentration zero. That is 99% of blanks should be less than  $L_c$ . Therefore less than 1% of method blanks should equal or exceed  $L_c$  (e.g., on average 99% of method blanks would be 'ND' using  $L_c$  as the censoring point). Use all applicable method blanks in the evaluation. Applicable method blanks are, at a minimum, all method blanks from analytical batches to be used to report results.

5.1.2 Monitor routine method blanks for positive results at or above the  $L_c$  on an ongoing basis. When a positive value at or above  $L_c$  is encountered, evaluate it for obvious errors. If the value is verified, either reject the batch or accept the batch and review the historic method blanks for other exceedences. If more than two of the last 100 method blanks have exceeded the  $L_c$  then either reject the method blank and batch and initiate corrective action/reanalysis, or adjust the  $L_c$  upward one digit (e.g., from 0.2 to 0.3) and reassess the exceedences. If adjusting the  $L_c$  by one digit does not bring the blank into control or if any method blank is above QL, then either the level or variability of contamination has increased or the calibration is biased at low concentrations. Either initiate corrective actions (e.g., recalibrate) or establish a new  $L_c$  in accord with Section 4.2.

5.1.3 On a minimum of an annual basis, review the method blank QC practices for compliance and document the review. For methods that produce at least some uncensored method blank data for an analyte, if the last 300 method blanks did not produce a method blank above the established  $L_c$ , lower  $L_c$  by one digit (e.g. 0.2 to 0.1). Re-evaluate the method blanks as in 5.1.2 using this lower  $L_c$ .

5.1.4 Where for an entire year, no uncensored method blank data were produced, use the QLC data (section 5.2.2) to determine the minimum area counts/other units of uncalibrated detector response to determine the response equal to the concentration of the  $L_c$ . Examine the user accessible setting and adjust them such that these set points censor response no higher than one-half of the  $L_c$ . Also, use the QLC to verify the QL is at least two times the  $L_c$  for these fully blank-censored analytes.

### 5.2 Ongoing verification of QL using the QLC

---

- 5.2.1 When the 100 QLC results have been collected, control chart limits must be calculated as follows:

$$\bar{r} \pm 3 \times s_r$$

$\bar{r}$  = mean spike recovery

$s_r$  = standard deviation of the set of spike recoveries

$$r = (y / QL) \times 100$$

y = measured concentration for a replicate sample spiked at QL

QL = known concentration of the quantitation limit check sample

Note that the acceptance criteria for bias and precision in Sections 1 and 2 must also be satisfied:

$$(s_r / \bar{r}) \times 100 \leq 20\% \text{ (equivalently, } \bar{r} \geq 5 \times s_r \text{ )}$$

$$s_r \leq 20\%$$

$$50\% \leq \bar{r} \leq 150$$

- 5.2.2 The laboratory must routinely analyze at least one QLC per batch of samples. For multi-analyte methods, which require multiple QLCs, the QLCs may be rotated on a batch to batch basis over a maximum of three batches. The recoveries of all QLC samples must fall within the control chart limits calculated in Section 5.2.1. When multiple analytes are simultaneously analyzed; if there are less than 20 analytes no more than 10% may fall between three and four standard deviations of the mean recovery. If there are 20 or more analytes, no more than 5% of the analytes may fall between three and four standard deviations of the mean recovery. The laboratory must report non-detects to values no less than the QL unless method performance is demonstrated at lower concentrations (e.g., via the use of lower spike concentrations for the quantitation limit check sample).
- 5.2.3 On an ongoing basis, evaluate the results from the QLCs. Verify that the recoveries fall within the calculated control chart limits, with no more than 5% of all QLC recoveries of each analyte falling outside four standard deviations of the mean recovery. If more than 5% of all QLC recoveries for any analyte falls outside four standard deviations of the mean recovery calculate a new estimate of QL per Section 4.2 and new control limits as discussed in Section 5.2.1 using the most recent 100 results.
- 5.2.4 When over 100 QLC results have been obtained, verify that the rate of false negatives and zero percent recoveries is no more than 1%. If this is not verified, a change in sensitivity may have occurred (e.g., the estimate of QL may be set too low). Calculate a new estimate of QL per Section 4.2 and new control limits as discussed in Section 5.2.1 using the most recent 100 results.

- 5.2.5 Replace the previously established QL with new QL or optionally, compare the variance of the new set of method blanks or quantitation limit check samples with that of the original (the initial demonstration of performance) using a two-tailed F test

$$F = \frac{(s_H^2)}{(s_L^2)}$$

$s_H^2$  = Larger variance estimate

$s_L^2$  = Smaller variance estimate

If the calculated value of F is greater than the F-statistic value for the 95% or 99% level of confidence, then the critical value and detection limit must be updated. (The 99% level of confidence may be used when a large number of analytes is being simultaneously determined.) If the calculated value of F is less than the F-statistic value for the 95% or 99% level of confidence, then the data should be pooled to calculate a more reliable estimate of  $L_C$ . However, it should be noted that the F-test is not robust to departures from normality (e.g., the underlying assumption for the calculation of  $L_C$ ).<sup>10</sup>

- 5.3 Whenever major method or instrument changes have occurred, perform the false positive verification test as discussed in section 4.1.2.4. If significant changes in the sensitivity estimate are indicated, repeat the initial detection estimate procedure.

## 6.0 DATA REPORTING

- 6.1 List the quantitation limit QL on reports. Detections may be reported to the critical value  $L_C$ . If detections less than the quantitation limit, QL, are required to be reported, then detections greater than or equal to the  $L_C$  but less than QL must be reported as quantitatively estimated values. Report measurements less than  $L_C$  as "< QL" (e.g., or "QL U," where the "U" qualifier indicates a non-detect). For methods capable of reporting uncensored results, users may request that results less than  $L_C$  be reported (for example, when statistical evaluations will be performed on the data).<sup>11</sup> Reported results less than  $L_C$  must be clearly identified as such. One example of appropriate reporting would be to report as "< QL (Y)", where Y denotes the numerical result obtained from the analytical method.
- 6.2 Quantitation limits should be reported to two significant figures and critical values should be reported using only one significant figure. When using single results for regulatory compliance, results below QL should be reported as Detected but Not Quantified (DNQ) and results below  $L_C$  should be reported as Not Detected. For example, if  $L_C = 0.6$ , and the QL, = 2.0; then results are reported as follows (a "J flag" is applied to denote quantitatively estimated results):

| Instrument Result | Reported Result              |
|-------------------|------------------------------|
| 2.1               | 2.1                          |
| 1.9               | DNQ or 2 J                   |
| 0.92              | DNQ or 0.9 J                 |
| 0.64              | DNQ or 0.6 J                 |
| 0.38              | ND or < 2 or 2U or <2 (0.38) |

In this example, the "Limit of Quantitation" (QL) is defined as the lowest concentration that, in the context of some tolerance for uncertainty, produces quantitatively reliable results for the end use of the data. (QL= 2 in this example, for the purposes of illustration only).

## **7.0 DEMONSTRATION OF SENSITIVITY IN VARIOUS MATRICES OR WHEN A SEQUENCE OF NON-ROUTINE STEPS ARE USED**

7.1 For some applications it may not be necessary to derive a Critical Value and Quantitation Limit for each matrix or when a sequence of non-routine steps are used. In these situations you should demonstrate whether or not  $L_C$  and QL can be achieved. Examples of this situation include:

- A direct request by a client to demonstrate the  $L_C$  and QL can be achieved in their matrix.
- When a new matrix with no past history yields non-detects.
- A sample stream with a history of detects near the  $L_C$  or QL starts giving non-detects.
- When non-routine steps are added to a method for which  $L_C$  and QL are already established.
- When routine steps are omitted from a method for which  $L_C$  and QL are already established.

7.1.1 If the effects of these non-routine conditions are of concern for samples with concentrations near quantitation, fortify a field sample as a matrix spike at the QL (QLMS).

7.1.2 If precision is also a concern, prepare a second spike at the QL on a split of the same field sample (QLMSD).

7.1.3 Process the QLMS and QLMSD through all preparatory and analytical steps.

7.1.4 Calculate the recoveries (and precision if QLMSD).

7.1.5 If the recovery and precision is outside of the control limits specified in Section 4.2.6 there is a presumed matrix effect and guidance in 7.2 should be used.

## 7.2 Determination of a Matrix Specific or Non-Routine Procedure specific $L_C$ and QL

For some applications, when recovery or bias of the QLMS/QLMSD in section 7.1 fall outside the control limits, an estimate of the  $L_C$  and QL in that specific matrix may be desired. Section 4.1 may be followed in its entirety for the matrix or non-routine procedure of interest. However, in some situations the minimum requirement of seven replicates may be too onerous. Examples of this situation include:

- When the required reporting level is greater than the laboratory quantitation limit
- When a non-routine analyte is added to a method at the specific request of a client and a higher degree of uncertainty can be tolerated (relative to the degree of acceptable uncertainty for the analytes routinely reported).
- When a client-specific matrix must be evaluated to determine, whether or not the sensitivity of the method in this matrix is adequate or substantially different than the sensitivity in a "clean" matrix such as reagent water.

7.2.1 For these situations analyze a minimum of four method blanks (for uncensored methods) or matrix samples from the same source spiked at or below the level at which sensitivity demonstration is required. (The matrix may be a reference matrix or a matrix applicable to a particular site or project). All blank and spike results must be used for next step.

7.2.2 Calculate the mean and standard deviation of the results. Determine the level of the analyte(s) in method blanks (*e.g.* the mean). Compare the mean and standard deviation with acceptance criteria in the method. If no acceptance criteria are available then recovery must be greater than or equal to 40% and the relative standard deviation must be less than or equal to 50% and the levels in method blanks must be below the required reporting limit or a full critical level and quantitation limit determination according to sections 4.1.1 or 4.1.2 must be performed.

- 7.2.3 For discerning differences in sensitivity between various matrices the F test may be used as outlined in 4.2.4.<sup>5</sup>
- 7.2.4 Gather additional data as the method is performed and calculate the critical level and quantitation limit according to sections 4.1.1 or 4.1.2 once sufficient data is available.
- 

## 8.0 ACKNOWLEDGEMENTS

8.1 The following members of Committee One on Detection contributed to the development of this procedure. Each member represents one or more of the Consensus organizations.

- Steven Bonde – Pacific Northwest Laboratories
- ✓ Richard Burrows, Ph.D. ([rburrows@stl-inc.com](mailto:rburrows@stl-inc.com)) – American Council of Independent Laboratories (ACIL), Severn Trent Laboratories (STL)
- Roger Claff, P.E. – American Petroleum Institute (API)
- Nancy Grams – Advanced Earth Technologies (AET), American Society for Testing and Materials D19 (ASTM), Waste Management
- ✓ Thomas Georgian, Ph.D. ([Thomas.Georgian@nwd02.usace.army.mil](mailto:Thomas.Georgian@nwd02.usace.army.mil)) – United States Army Corps of Engineers (USACE), Standard Methods
- Angie Grooms – Duke Power, Utility Water Act Group (UWAG)
- Donna Hill – Southern Company, Utility Water Act Group (UWAG)
- Larry LaFleur – American Forestry and Paper Association, Inter-Industry Analytical Group (IIAG)
- Patty Lee – Hampton Road Sanitary District (HRSD), Water Environment Federation (WEF)
- Judith Morgan – Environmental Science Corporation
- ✓ Kenneth Osborn ([kosborn@ebmud.com](mailto:kosborn@ebmud.com)) – East Bay Municipal Utility District (EBMUD), American Water Works Association (AWWA), Standard Methods
- John Phillips ([jphill12@ford.com](mailto:jphill12@ford.com)) – Ford Motor Company, Alliance of Automobile Manufactures (AAM), Inter-Industry Analytical Group (IIAG), Michigan Environmental Laboratories Association (MELA)
- Jim Pletl – Hampton Road Sanitary District (HRSD), Association of Metropolitan Sewerage Agencies (AMSA)
- Lennie Routten – Northrop Grumman, VA American Water Works Association (AWWA), VA Water Environment Association (VWEA)

- Rock Vitale - Environmental Standards, Inc., representing Environmental Laboratory Advisory Board (ELAB), American Institute of Chemists (AIC), National Registry of Certified Chemists (NRCC)

➤ *Indicates committee chairman*

✓ *Indicates a major contributor who can be contacted to answer technical questions*

## **9.0 ATTACHMENTS**

9.1 Spreadsheet

9.2 Flowchart

9.3 Grubbs Test

Table 1: K Values

| <b>v</b> | <b><math>K_{0.99, 0.01, v}</math></b> |
|----------|---------------------------------------|
| 6        | 6.101                                 |
| 7        | 5.529                                 |
| 8        | 5.127                                 |
| 9        | 4.829                                 |
| 10       | 4.599                                 |
| 11       | 4.415                                 |
| 12       | 4.264                                 |
| 13       | 4.138                                 |
| 14       | 4.031                                 |
| 15       | 3.939                                 |
| 16       | 3.859                                 |
| 17       | 3.789                                 |
| 18       | 3.726                                 |
| 19       | 3.67                                  |
| 20       | 3.619                                 |
| 21       | 3.573                                 |
| 22       | 3.532                                 |
| 23       | 3.494                                 |
| 24       | 3.458                                 |
| 25       | 3.426                                 |
| 26       | 3.396                                 |
| 27       | 3.368                                 |
| 28       | 3.342                                 |
| 29       | 3.317                                 |
| 30       | 3.295                                 |
| 31       | 3.273                                 |
| 34       | 3.216                                 |
| 35       | 3.199                                 |
| 36       | 3.182                                 |
| 37       | 3.167                                 |
| 38       | 3.152                                 |
| 39       | 3.138                                 |
| 40       | 3.125                                 |
| 41       | 3.112                                 |
| 42       | 3.100                                 |
| 43       | 3.088                                 |
| 44       | 3.077                                 |
| 45       | 3.066                                 |
| 46       | 3.055                                 |

47 3.045  
48 3.036  
49 3.027  
50 3.018  
51 3.009  
52 3.001  
53 2.993  
54 2.985  
55 2.977  
56 2.97  
57 2.963  
58 2.956  
59 2.949  
60 2.943  
61 2.936  
62 2.93  
63 2.924  
64 2.919  
65 2.913  
66 2.907  
67 2.902  
68 2.897  
69 2.892  
70 2.887  
71 2.882  
72 2.877  
73 2.873  
74 2.868  
75 2.864  
76 2.86

78 2.851  
79 2.847  
80 2.843  
81 2.839  
82 2.836  
83 2.832  
84 2.828  
85 2.825  
86 2.821  
87 2.818  
88 2.815  
89 2.811  
90 2.808

|    |       |
|----|-------|
| 91 | 2.805 |
| 92 | 2.802 |
| 93 | 2.799 |
| 94 | 2.796 |
| 95 | 2.793 |
| 96 | 2.79  |
| 97 | 2.787 |
| 98 | 2.784 |
| 99 | 2.782 |

For large n (e.g., n > 100),

$$K_{0.99,0.01,v} = Z_{0.99} \sqrt{v/\chi_{0.01}^2} \approx Z_{0.99} \sqrt{\frac{v}{(Z_{0.01}(2v)^{1/2} + v)}} = 2.326 \sqrt{\frac{v}{v - 2.326(2v)^{1/2}}}$$

Table 2: Student t values

| <b>v</b> | <b><math>t_{0.995,v}</math></b> |
|----------|---------------------------------|
| 6        | 3.707                           |
| 7        | 3.499                           |
| 8        | 3.355                           |
| 9        | 3.250                           |
| 10       | 3.169                           |
| 11       | 3.106                           |
| 12       | 3.055                           |
| 13       | 3.012                           |
| 14       | 2.977                           |
| 15       | 2.947                           |
| 16       | 2.921                           |
| 17       | 2.898                           |
| 18       | 2.878                           |
| 19       | 2.861                           |
| 20       | 2.845                           |
| 21       | 2.831                           |
| 22       | 2.819                           |
| 23       | 2.807                           |
| 24       | 2.797                           |
| 25       | 2.787                           |
| 26       | 2.779                           |
| 27       | 2.771                           |
| 28       | 2.763                           |
| 29       | 2.756                           |
| 30       | 2.750                           |
| 31       | 2.744                           |
| 32       | 2.738                           |
| 33       | 2.733                           |
| 34       | 2.728                           |
| 35       | 2.724                           |
| 36       | 2.719                           |
| 37       | 2.715                           |
| 38       | 2.712                           |
| 39       | 2.708                           |
| 40       | 2.704                           |
| 50       | 2.678                           |
| 60       | 2.660                           |
| 70       | 2.648                           |
| 80       | 2.639                           |
| 90       | 2.632                           |
| 100      | 2.626                           |
| 150      | 2.609                           |
| 200      | 2.601                           |

Infinite 2.576

<sup>1</sup> If the number of replicates is small, a more reliable estimate of the critical value can be made using the upper tolerance interval (Georgian and Osborn, Quality Assurance, 8:1-9, 2003).

Unless a large number of replicates is available,  $L_C$  is calculated a formula of the form:

$$L_C = z_p s_{UCL, 1-\gamma} = z_p \sqrt{v/\chi_{v, \gamma}^2} s = K_{p, \gamma, v} s$$

where  $\chi_{v, \gamma}^2$  is the  $\gamma$ 100<sup>th</sup> percentile of the  $\chi^2$  distribution with  $v$  degrees of freedom and  $z_p$  denotes the  $p$ 100<sup>th</sup> percentile of the standard normal distribution. Note that normality is being assumed. This equation is the  $(1 - \gamma)$  100% tolerance interval that contains at least the proportion  $p$  of the population. If a large number of replicate measurements were to be performed for a blank, then  $p$ 100% of the measurements would be less than the critical value  $L_C$  with  $(1 - \gamma)$  100% confidence. If  $p = 0.99$  and  $\gamma = 0.01$ , 99% percent of all future measurements will be less than  $L_C$  with 99% confidence. For  $p = 0.99$ ,  $\gamma = 0.01$ , and  $v = 6$  ( $n = 7$ ),

$$L_C = 2.33\sqrt{6/0.872} s \approx 6 s$$

<sup>2</sup> This is very similar to Currie's critical level,  $L_C$  (Anal. Chem. Vol. 40, No. 3 March 1968, p586). It is the level at which a result can be confidently distinguished from the blank. The detection limit,  $L_D$ , is set at the lowest value that can be reliably detected, where  $L_D \approx 2 L_C$ , assuming constant variance and no analytical bias. Currie's original procedure assumes that the mean result from the blanks will be subtracted from any individual blank result (or equivalently, the absence of significant analytical bias for the final reported result). Since blank subtraction is not allowed in most environmental methods, the mean of the blanks must be added into the calculation.

<sup>3</sup> If spikes at  $L_D$  have low average recovery (<70%), then  $L_D$  may need to be increased in order to maintain reliable detection (i.e., a result in a measured value greater than  $L_C$ ). The approximate spiking concentration for  $L_D$  can be estimated from the equation:

$$L_D = \bar{X} + 2s K_{p, \gamma, v} \left( \frac{100}{\% \text{ Recovery}} \right)$$

<sup>4</sup> If the sample mean ( $\bar{X}$ ) is not significantly different from zero, then the detection limit and critical value may also be determined using the same equation with  $\bar{X} = 0$ . It is recommended that a two-tailed t-test be performed at the 99% level of confidence to determine if the mean is significantly different from zero. To do this, calculate the following:

$$t = |\bar{X}| / (s/\sqrt{n})$$

Compare this to the critical value of the Student's t distribution, which is denoted by  $t_{1-\alpha/2, v}$ . For a set of  $n$  replicate measurements and the  $(1-\alpha)$ 100% level of confidence, this value is the  $(1-\alpha/2)$ 100<sup>th</sup> percentile of the Student's t distribution with  $v = n - 1$  degrees of freedom. Table 2 lists the critical values for the Student's t-distribution for the 99% level of confidence (for a two-tailed test) for various degrees of freedom. For  $n = 7$  and  $\alpha = 0.01$ ,  $t_{0.995, 6} = 3.71$ . If  $-t_{1-\alpha/2, v} < t < t_{1-\alpha/2, v}$ , then there is insufficient evidence to conclude (at the specified level of confidence) that the "true" (population) mean is significantly different from zero and the detection limit and critical value may be calculated by setting the mean equal to zero. The mean is significantly different from zero at the  $(1-\alpha)$ 100% level of confidence, if  $|t| > t_{1-\alpha/2, v}$ .

<sup>5</sup> Alternatively, a two-tailed F-test at the 95% or 99% level of confidence may be performed using the highest and the lowest standard deviation to determine if the standard deviations are significantly different from one another.

$$F = \frac{(s_H^2)}{(s_L^2)}$$

$s_H^2$  = Largest variance estimate

$s_L^2$  = Smallest variance estimate

If the calculated value F (above) is less than the critical value of this statistic for the  $(1-\alpha)100\% = 95\%$  or  $99\%$  level of confidence,  $F(n_H-1, n_L-1, 1-\alpha/2)$ , then it may be concluded that the instruments possess similar sensitivities and the pooled standard deviation may be used. The values  $n_H$  and  $n_L$  denote the number of replicates used to calculate the standard deviations  $s_H$  and  $s_L$ , respectively. If seven replicates are used for all instruments, then the critical values for the 95% and 99% confidence levels are  $F(6, 6, 0.975) = 5.82$  and  $F(6, 6, 0.995) = 11.07$ , respectively. Note that the square root of the critical value is about 2.4 for the 95% level of confidence and 3.3 for the 99% level of confidence. Therefore, if the instruments possess similar sensitivities, the largest standard deviation should not be greater than approximately two or three times the smallest standard deviation. The 95% level of confidence should be used when a small number of analytes are being simultaneously determined (e.g.,  $n < 10$ ); the 99% level of confidence is recommended when a large number of analytes is being determined (e.g.,  $n > 10$ ).

<sup>6</sup> An apparent minimum signal to noise ratio of approximately 3:1 is a reasonable way to evaluate if a measurable signal is achievable for a chromatographic determination. This evaluation need not be done quantitatively, but may be simply estimated visually. For this evaluation integration conditions must be identical to routine operating conditions (e.g. no modification of integration threshold values, peak slope, etc.).

It is critical that the qualitative identification procedures that are used be the same as those employed during routine analysis. Consider the example of a GC/MS test that requires qualifier ions be present in a specific ratio range relative to the primary ion and the identification is rejected if they are not present in that range. In this case it is not acceptable to call the analyte detected for the replicate detection limit study if the qualifier ratios fail. In general, the instrument's target compound identification software should automatically identify the peak without manual intervention to establish reliable qualitative identification. In the case of dual column confirmatory analysis the analyte must be detected on both columns. In the case of pattern recognition, for example Aroclor analysis, the analyst must be able to recognize the pattern.

<sup>7</sup> It is recognized that  $x$  will probably not be the same for each analyte in a multi-analyte test.

<sup>8</sup> For some isotope dilution methods an Estimated Detection Limit (EDL) may be measured and calculated for each analytical result. The EDL reflects the sample concentration of the particular analyte which would be required to cause a peak at 2.5 times the background noise (i.e. a positive result) for the particular analysis. The EDL is calculated according to the following equation:

$$\text{Estimated Detection Limit} = \frac{N \times 2.5 \times Q_{is}}{H_{is} \times RRF \times W \times S}$$

where:

- N = peak to peak noise of quantitation ion signal in the region of the ion chromatogram where the compound of interest is expected to elute
- H<sub>is</sub> = peak height of quantitation ion for appropriate internal standard
- Q<sub>is</sub> = ng of internal standard added to sample
- RRF = mean relative response factor of compound obtained during initial calibration
- W = amount of sample extracted (grams or liters)
- S = percent solids (optional, if results are requested to be reported on dry weight basis)

If this procedure is used, determination of the detection limit is the higher of the EDL or  $L_d$ .

<sup>9</sup> Where the population sample has more than one method blank value exceeding  $L_c$ ,  $L_c$  is suspect. Where more than two method blanks of the last 100 exceed  $L_c$  the laboratory may either increase  $L_c$  by one unit (e.g., 0.04 to 0.05) or it may repeat the estimation procedure. Continue to monitor and adjust  $L_c$  as needed.

<sup>10</sup> Due to departures from normality, it could be concluded that the variances are significantly different when they are not. Therefore, it is strongly recommended that Levene's test be used in lieu of the F test to determine if the standard deviations (variances) differ at 95% or 99% level of confidence.

<sup>11</sup> The International Union of Pure and Applied Chemistry (IUPAC) recommendation is to "...*always* report both the estimated value of the measured quantity ( $\hat{L}$ ) and its uncertainty, *even* when  $\hat{L} < L_c$  results in the decision 'not detected.' Otherwise, there is needless information loss, and, of course, the impossibility of averaging a series of results." (*Chemometrics and intelligent laboratory systems* (Currie, 1997), p. 156). Reporting below  $L_c$  may not be practical, however, and reported values below  $L_d$  would be inappropriate for making compliance determinations. However, measurement values below  $L_c$  can be used in statistical evaluations of large data sets for the purpose of calculating means, standard deviations, confidence intervals, and future event likelihoods. Examples of such use would include the comparison of two analytical methods in the region of analytical lower limit capability, determination of a background contribution allowance, and reasonable potential analysis.