Designation: ~~D7783 – 13~~ D7783 – 21

Standard Practice for Within-laboratory Quantitation Estimation (WQE)[1]

This standard is issued under the fixed designation D7783; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

~~**Note—Balloted information was included and the year date changed on March 28, 2013.**~~

1. Scope

1.1 This practice establishes a uniform standard for computing the within-laboratory quantitation estimate associated with *Z* % relative standard deviation (referred to herein as $WQE_{~~Z~~ Z\ \%}$), and provides guidance concerning the appropriate use and application.

1.2 $WQE_{~~Z~~ Z\ \%}$ is computed to be the lowest concentration for which a single measurement from the laboratory will have an estimated ~~*Z* %~~*Z* % relative standard deviation (~~*Z* %~~(*Z* % RSD, based on within-laboratory standard deviation), where *Z* is typically an integer multiple of 10, such as 10, 20, or 30. *Z* can be less than 10 but not more than 30. The $WQE_{10\ \%}$ is consistent with the quantitation approaches of Currie (**1**)[2] and Oppenheimer, et ~~a.~~al. (**2**).

1.3 The fundamental assumption of the WQE is that the media tested, the concentrations tested, and the protocol followed in ~~the~~ developing the study data provide a representative and fair evaluation of the scope and applicability of the test method, as written. Properly applied, the WQE procedure ensures that the WQE value has the following properties:

1.3.1 *Routinely Achievable WQE Value*—The laboratory should be able to attain the WQE in routine analyses, using the ~~laboratory's~~laboratory's standard measurement system(s), at reasonable cost. This property is needed for a quantitation limit to be feasible in practical situations. Representative data must be used in the calculation of the WQE.

1.3.2 *Accounting for Routine Sources of Error*—The WQE should realistically include sources of bias and variation that are common to the measurement process and the measured materials. These sources include, but are not limited to intrinsic instrument noise, some typical amount of carryover error, bottling, preservation, sample handling and storage, analysts, sample preparation, instruments, and matrix.

1.3.3 *Avoidable Sources of Error Excluded*—The WQE should realistically exclude avoidable sources of bias and variation (that is, those sources that can reasonably be avoided in routine sample measurements). Avoidable sources ~~would~~ include, but are not limited to, modifications to the sample, modifications to the measurement procedure, modifications to the measurement equipment of the validated method, and gross and easily discernible transcription errors (provided there ~~was~~is a way to detect and either correct or eliminate these errors in routine processing of samples).

1.4 The WQE applies to measurement methods for which instrument calibration error is minor relative to other sources, because

[1] This practice is under the jurisdiction of ASTM Committee D19 on Water and is the direct responsibility of Subcommittee D19.02 on Quality Systems, Specification, and Statistics.

Current edition approved ~~March 28, 2013~~Nov. 15, 2021. Published ~~April 2013~~March 2022. Originally approved in 2012. Last previous edition approved in ~~2012~~2013 as ~~D7783 – 12.~~D7783 – 13. DOI: ~~10.1520/D7783-13.~~10.1520/D7783-21.

[2] The boldface numbers in parentheses refer to ~~a~~the list of references at the end of this standard.

this practice does not model or account for instrument calibration error, as is true of ~~quantiation~~ most quantitation estimates in general. Therefore, the WQE procedure is appropriate when the dominant source of variation is not instrument calibration, but is perhaps one or more of the following:

1.4.1 *Sample Preparation,* and especially when calibration standards do not go through sample preparation.

1.4.2 *Differences in Analysts,* and especially when analysts have little opportunity to affect instrument calibration results (as is the case with automated calibration).

1.4.3 *Differences in Instruments (measurement equipment),* such as differences in manufacturer, model, hardware, electronics, sampling rate, chemical-processing rate, integration time, software algorithms, internal signal processing and thresholds, effective sample volume, and contamination level.

1.5 *Data Quality Objectives*—For a given method, one typically would compute the ~~lowest % RSD possible for any given data set.~~ WQE for the lowest RSD for which the data set produces a reliable estimate. Thus, if possible, $WQE_{10\ \%}$ would be computed. If the data indicated that the method was too noisy, so that $WQE_{10\ \%}$ could not be computed reliably, one might have to compute instead $WQE_{20\ \%}$, or possibly $WQE_{30\ \%}$. In any case, a WQE with a ~~higher % RSD~~ higher RSD level (such as $WQE_{50\ \%}$) would not be considered, though a WQE with RSD ~~<10 %~~ < 10 % (such as $WQE_{~~1 \%~~5\ \%}$) ~~would~~could be acceptable. The appropriate level of ~~% RSD~~ RSD is based on the ~~data-quality~~ data quality objective(s) for a particular use or uses. This practice allows for calculation of WQEs with user selected ~~% RSDs~~ RSDs less than 30 %.

1.6 *This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.*

2. Referenced Documents

2.1 *ASTM Standards:*[3]

D1129 Terminology Relating to Water

D2777 Practice for Determination of Precision and Bias of Applicable Test Methods of Committee D19 on Water

~~D6091~~E2586 Practice for ~~99 %/95 % Interlaboratory Detection Estimate (IDE) for Analytical Methods with Negligible Calibration Error~~Calculating and Using Basic Statistics

~~D6512 Practice for Interlaboratory Quantitation Estimate~~

~~D7510 Practice for Performing Detection and Quantitation Estimation and Data Assessment Utilizing DQCALC Software, based on ASTM Practices D6091 and D6512 of Committee D19 on Water~~

~~E1763 Guide for Interpretation and Use of Results from Interlaboratory Testing of Chemical Analysis Methods~~

2.2 *BIPM Documents:*

GUM: JCGM 100:2008 Evaluation of measurement data—Guide to the expression of uncertainty in measurement[4]

3. Terminology

3.1 *Definitions*—For definitions of terms used in this practice, refer to Terminology D1129~~.~~, Practice E2586, and the GUM.

3.2 *Definitions of Terms Specific to This Standard:*

3.2.1 *censored measurement, n*—a measurement that is not reported numerically, but is stated as a "nondetection" or a less-than (for example, "less than 0.1 ppb").

3.2.2 *quantitation limit (QL) or limit of quantitation (LQ), n*—a numerical value, expressed in physical units or proportion, intended to represent the lowest level of quantitation, based on a set of criteria for quantitation.

3.2.2.1 *Discussion*—

The WQE is an example of a ~~QL~~QL.

3.2.3 *Z % within-laboratory quantitation estimate*

[3] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

[4] Available from https://www.bipm.org/en/publications/guides/, accessed May 2021.

(WQE$_{Z\ \%Z\ \%}$), n—(in accordance with Currie (**1**))—The lowest concentration for which a single measurement from the examined laboratory will have an estimated ~~Z %~~Z % relative standard deviation (~~Z %~~(Z % RSD, based on the within-laboratory standard deviation).

4. Summary of Practices

4.1 The WQE procedure provides an estimate of the true concentration at which a desired ~~level of (relative)~~ relative precision is achieved. Whether from analysis of routine quality samples or from studies undertaken from time to time (or both), the first step is to acquire data representative of the laboratory performance for use in the WQE calculations. Such data must include concentrations suitable for modeling the precision and bias over a range of concentrations. Each datum for a method/matrix/analyte should represent an independent sample where routine sources of measurement variability occur at typical levels of influence. Outlying individual measurements should be eliminated, using an accepted, scientifically-based procedure for outlier identification and a documented, scientific basis for removal of data from the data set, such as found in Practice D2777. WQE computations must be based on retained data (after optional outlier removal) from at least six independent measurements at a minimum of five concentrations.

4.2 Retained data are analyzed to identify and fit one of four proposed standard-deviation ~~models.~~models: *constant,linear* (straight-line), *hybrid* (proposed by Rocke and Lorenzato (**3**)), and *exponential.* These models describe the relationship between the within-laboratory standard deviation of measurements and the true concentration, *T*. The ~~identification~~selection process involves evaluating the ~~models in order, from simplest to most complex: constant, straight-line, exponential, and hybrid (proposed by Rocke and Lorenzato~~models, starting with the linear model and performing statistical tests to choose the (simplest~~3~~) ~~and Guide~~model that E1763. adequately fits the data. Evaluation includes ~~statistical-significance~~ statistical significance testing and residual analysis, and ~~is based on the best~~ requires the judgment of a qualified ~~chemist and the requirement to utilize the simplest model that adequately fits the data.~~chemist.

4.3 Once the standard-deviation model has been ~~determined, it is used to determine~~ selected, it determines the fitting technique for ~~modeling~~ the model of measured concentration ~~(referred~~ versus true concentration, referred to in this practice as the *mean-recovery ~~model)~~model.* ~~to true concentration. If~~ If the standard deviation is constant, then ordinary least squares ~~is~~ may be used. If the standard deviation is not constant, the ~~modeled-standard-deviation predictions~~predicted standard deviations are used to generate weights for use in ~~the weighted-least-squares fitting. With either fitting technique, a straight line is the model that is fitted~~ weighted least squares. Regardless of the fitting technique, the mean-recovery model fits a straight line to the data.

4.4 The ~~liner~~linear regression ~~(true~~(measured versus ~~measured)~~true) is evaluated for statistical significance, for lack of fit, and for residual patterns.

4.5 These two models ~~(standard-deviation~~ (standard deviation and ~~calibration)~~recovery) are then used to calculate the WQE values. Either a direct calculation or ~~interactive~~iterative algorithm (depending on the model) is used to compute WQE$_{10\ \%}$, the lowest true concentration with estimated RSD = 10 % (~~Z~~(*Z* = 10); WQE$_{20\ \%}$ ~~(% RSD=20 %=Z);~~ (%RSD = 20 = *Z*); and WQE$_{30\ \%}$ ~~30 % (% RSD=30 %=Z).~~ (%RSD = 30 = *Z*). If needed for particular data-quality objectives (DQOs), WQE$_{Z\ \%Z\ \%}$ may be computed for some $Z < 10$. The particular ~~Z %~~Z % selected for use should depend upon the data-quality needs and the realized performance. Typically, either 10 % or 20 % is used in ~~environmental-water~~ environmental water testing. The 30 % RSD approaches the criterion for detection. *Z* values greater than 30 should not be used. An RSD of 5 % approximates a level at which at least one sure significant digit has been achieved.

5. Significance and Use

5.1 Appropriate application of this practice should result in a WQE achievable by the laboratory in applying the tested method/matrix/analyte combination to routine sample analysis. That is, a laboratory should be capable of measuring concentrations greater than WQE$_{Z\ \%Z\ \%}$, with the associated RSD equal to ~~Z %~~Z % or less.

5.2 The WQE values may be used to compare the quantitation capability of different methods for analysis of the same analyte in the same matrix within the same laboratory.

5.3 The WQE procedure should be used to establish the within-laboratory quantitation capability for any application of a method in the laboratory where quantitation is important to data use. The intent of the WQE is not to impose reporting limits. The intent is to provide a reliable procedure for establishing the quantitative characteristics of the method (as implemented in the laboratory

for the matrix and analyte) and thus to provide the laboratory with reliable information characterizing the uncertainty in any data produced. Then the laboratory ~~may~~can make informed decisions about censoring data and has the information necessary for providing reliable estimates of uncertainty with reported data.

6. Procedure

6.1 This procedure is described in stages as follows: Development of Data, Data Screening, Modeling Standard Deviation, Fitting the Recovery Relationship, and Computing the Quantitation Estimates.

6.2 *Development of Data for Input to the Calculations*—A single WQE calculation is performed per analyte, matrix/medium and method. A minimum of five concentrations must be used to allow for high-quality estimation of ~~true-verses-measured~~measured-versus-true concentration, and for modeling the relationship of standard deviation to true concentration. ~~A minimum of~~ At least six values at each concentration are required to provide a high-quality estimation of the standard-deviation and the recovery relationships. Additional concentrations (especially ~~additional,~~additional representative, independent samples at each concentration) are highly ~~encouraged; such~~encouraged. Such inclusion will reduce the uncertainty in the estimate and better assure that after outlier removal, the minimum requirements for concentrations and values will be met. Data for each WQE calculation should come from only one ~~laboratory,~~ laboratory and one method, and be for only one analyte in one matrix/medium. Concentrations may be designed in advance, or data already developed may be used. ~~For multi-laboratory determinations, see Practice D6091.~~

6.2.1 *Designing Concentrations*—Where concentrations are being selected in advance of the collection of data, the development of an optimized design should consider many factors, including:

6.2.1.1 Concentrations of available data, such as routine quality-control samples.

6.2.1.2 Potential use of the same data to calculate detection limits and or other control limits.

6.2.1.3 The anticipated or previously determined WQE (study range should exceed this value by at least a factor of 2).

6.2.1.4 The potential need to eliminate the lowest concentration(s) selected (see zero-concentration discussion above).

6.2.1.5 Where possible, select a WQE study design that has enough distinct concentrations to assess statistical lack of fit of the models (see Draper and Smith (**4**)). Recommended designs are: (*a*) The semi-geometric design with five or more true concentrations, T_1, T_2, and so forth, such as: 0, WQE_0, ~~D~~/D^2, WQE_0, ~~D,~~/D, WQE_0, $D \times WQE_0$, $D^2 \times WQE_0$, where D is a number greater than or equal to 2 and WQE_0 is an initial estimate of the WQE, (*b*) equi-spaced design: 0, $WQE_0/2$, WQE_0, $(3/2) \times WQE_0$, $2 \times WQE_0$, $(5/2) \times WQE_0$. Other designs with at least five concentrations—provided the design includes blanks, one concentration that approximates $2 \times WQE_0$, and at least one nonzero concentration below WQE_0—should be adequate.

6.2.2 *Considerations for All Concentration Selections:*

6.2.2.1 The range of the data, the number of unique concentrations, and the spacing of the ~~concentration~~concentrations are the primary decisions for study design, in addition to the number of replicates at each concentration. The range chosen, excluding the zero value for purposes of the discussion of range, should ~~be~~extend from below the estimated detection level to above the WQE of interest (for example, 10 %, 20 %, or 30 %), ~~so as to allow for performance of calculations without~~ to avoid the need for extrapolation.

6.2.2.2 A single model (one of the four models in this practice) should describe the behavior of the standard deviation in this range. The anticipated form of the relationship between measurement standard deviation and true concentration, if known, can help in choosing design spacing. Chemistry, physics, empirical evidence, or informed judgment may make one model more likely than others. Evaluation of interlaboratory method-validation studies may also provide information about these relationships. If a model of standard deviation is likely to be one with curvature at lower concentrations (hybrid or ~~exponential)~~exponential), then a semi-geometric design is favored. If the likely relationship is constant or ~~straight-line,~~linear, then equidistant spacing might be favored.

6.2.2.3 ~~Additional~~ Inclusion of additional concentrations, beyond the minimum of five concentrations, is strongly recommended where knowledge of these relationships is unknown. Where more than one order of magnitude is covered in the range selected (per range definition in 6.2.2.1), it is recommended that four additional unique concentrations be added per additional order of magnitude greater than one.

(1) Where ongoing quality-control (QC) information is available and it indicates that precision is good at the concentration of this quality control ~~measure,~~measure (for example, 5 % RSD or ~~less,~~less at higher concentrations), then establishing the maximum concentration for the study at or below that concentration should be considered where the ~~%~~ RSD criterion for the WQE is higher (for example, a $WQE_{10\ \%}$ ~~10 %).~~).

(2) Where ongoing QC demonstrates a high ~~%~~ RSD (for example, above 30 %), several concentrations at and above the concentration of the QC sample should be included.

NOTE 1—Where more than five concentrations are available, determination of the WQE with and without the highest (and potentially the lowest) concentration(s) included can provide insight into the effects of the highest concentration(s) on the recovery relationship and the modeling of standard deviation. Calculation of the WQE values based on the most appropriate and applicable concentrations, so long as minimums are met, is allowed.

6.2.2.4 The minimum of six independent values at each concentration is required by this practice to provide a minimally acceptable data set for calculation of standard deviation at each concentration. Increasing the number of levels is desirable where project constraints allow. It is not required that the same number of replicates be used for each concentration; however, extreme differences (for example orders of magnitude) should be avoided.

6.2.2.5 Known, routine sources of measurement variability, consistent with those of routine analysis of samples, must have been in action at the time of the generation of the data to be used, if the WQE is to be used for characterizing routine performance. That is, in order for the WQE to represent routinely achieved quantitation, the data used for WQE calculation must be generated under routine analytical conditions. Representative within-laboratory variation can only be seen if the number of qualified analysts and qualified measurement systems in the laboratory are represented. The data used and the more combinations included, the less effect any specific bias in these pairings should have on the WQE estimate. Similarly, sample management (for example, holding time) and allowed variations in routine sample-processing procedures must be included. The time period spanned must allow routine, time-dependent sources of variation to affect the testing. This consideration should include factors such as the frequency of calibration of instruments, introduction of newly prepared or purchased standards, reagents and supplies, and sample-holding times. Historically, the failure to utilize representative data in determination of quantitation limits has been a primary component in over-statements of quality through quantitation-limit values and should be strictly avoided (that is, garbage in, garbage out). Ideally, each measurement would be a double-blind measurement made by a different analyst, using a different (qualified) measurement system on a different day. Optimally, data to be used should be either completely blind, or from known but completely routine, integrated testing (such as routine quality-control data). In any case, the goal is to minimize special treatment of the WQE test samples.

6.2.2.6 Where the WQE is meant to represent the best possible performance, and not routine performance, then optimized conditions for data generation would be appropriate. Similarly, if the performance of only a single process, instrument system, analyst, ~~etc.~~etc., is of interest, only the applicable variables should be included. It is the responsibility of the user of this practice

to assure that the appropriate data are utilized for the end use(s) of WQE. Where the end use is unknown, the data generator who is using the WQE needs to disclose the specific attributes of the data used in the calculation (as well as the ~~% RSD),~~ RSD), and thus of the WQE.

6.2.2.7 Where preexisting, routine-source data (for example, quality-control data) are used, care must be taken to assure that: (*1*) each data point represents a true and independent sampling of the population (as well as of the sample medium being examined, where applicable) and (*2*) all sample-processing steps and equipment (for example, bottles, preservatives, holding, preparation, cleanup) are represented. Also, "true" concentration levels must either be known (that is, true "spiked" concentration levels), or knowable, after the fact. A concentration is considered *known* if reference standards can be purchased or constructed, and *knowable* if an accurate determination can be made.

6.2.2.8 Transformation of other types of data (such as laboratory replicates, which under-represent the variability as compared to independent samples and usually do not have known true concentrations), using scientifically and statistically sound approaches is not prohibited by this practice. However, care must be taken and the validity of these transformations tested. It is also critical that any standards used to prepare study samples be completely independent of the standards used to calibrate the instrument.

6.2.2.9 Blank correction should not be performed, unless the method requires this correction to calculate result values.

6.2.3 *True-Concentration Zero (Blank) Data Discussion*—Where possible, it is preferable to include data from samples with true concentration of zero (for example, blanks). However, for many methods, it may not be possible to conduct an unbiased sampling of the zero (blank) concentration samples, since instruments and software systems routinely smooth electronic information (raw data) from the detector and through software settings that censor reported data. Through these automated processes, many testing instruments return to the operator a result value of "zero," when, if these processes had been turned off, a non-zero numeric result (positive or negative) would have been produced. These "false-zero" values adversely affect the use of the zero-concentration data in statistics and should not be used for WQE studies. Most chromatography systems (and many other types of computer-assisted instruments) have instrument set-points (such as (digital) bunch rate, slope sensitivity, and minimum area counts) that are operator-controllable. For purposes of this study, generating as much uncensored low-level data as practical is important and the presence of these processes as well as the setting of any operator-controllable setting should be evaluated.

NOTE 2—Qualitative criteria used by the method to identify and discriminate among analytes are separate criteria, and must be satisfied according to the method.

6.2.3.1 Once true-concentration-zero measurements have been generated, and prior to use, it is important to examine and evaluate these data. A graph of measured concentration by frequency of occurrence may be helpful. However, unless a fairly large sample size is represented (for example, ~~*n>20),n*~~ ≥ 20), the distribution may be distorted by the random nature of sampling alone. As a general rule, if there were no bias, then on average and over a large sampling, a truly uncensored set of zero-concentration (blank) data would have a mean of zero with approximately half of the results being negative values and half positive, and be Normally distributed. If some positive or negative bias were present, the percentages would shift. However, in general the frequency should be higher near the mean of the values and should decline as the concentrations move away from the mean, with approximately half of the non-mean data above and half below the mean.

(1) Blank data are considered suspect if: (*1*) there is no variation in these data, (*2*) there are an inordinate number of zero values (and no negative values) relative to the frequencies of positive values (6.2.3 above), (*3*) if there is a high frequency of the lowest value in the data set (for example, where minimum-peak-area rejection has been used) relative to the frequency of higher concentration values, and few or no lower values, or (*4*) a frequency graphic does not begin to approximate a bell curve (when there are 20 or more samples).

(2) If the distribution of the data is suspect, the literature, plus instrument-software and equipment manuals, should be consulted. These documents can provide an understanding of: (*1*) the theory of operation of the detection system, (*2*) the signal processing, calibration, etc., and (*3*) other aspects of the conversion of response to reported values. Judgment will be needed to determine whether to use some or all of the true-concentration-zero (blank) data, or to exclude the data from the calculations. In general, if less than 10 % of the zero-concentration data are: (*1*) censored, (*2*) suspect, or (*3*) false-zeros, then these "problem" data should be removed. Only the remaining blank data are used in the WQE calculations; there must be at least six replicates. Where the zero concentration is excluded or is not possible to obtain, it is important to include a true concentration as close as possible to zero in the study design.

(3) Where 75 % or less of the data are censored or smoothed, and there are at least six remaining values, it is reasonable to use statistical procedures to simulate the distribution that is missing or smoothed. Software procedures are commercially available. Additionally, procedures such as log-normal transformation may be used to accommodate data that are not normally distributed. The ~~presence of zero-concentration in the study data and in the WQE is not as critical as inclusion of such data in the WDE calculations. Therefore, the~~ decision about inclusion or exclusion of zero-concentration data in a WQE data set should weigh: (*1*)

the number of other concentrations available, (*2*) the range of the other concentrations, and (*3*) the risk of extrapolation of the WQE outside the data-set concentration range against the quality of the zero-concentration data.

6.2.3.2 *True Concentrations Near Zero*—As with concentration zero, true concentrations very near to zero may also have been censored, smoothed, and contain false-zeros. Examination of these very low concentrations, as above for zero concentration, is important. The likelihood of occurrence and the percentage of data affected decreases with increasing concentration.

6.3 *Data Screening, Outlier Identification, and Outlier Removal:*

6.3.1 Data that are to be the input to the WQE calculation should be screened for compliance with this practice's conditions, appropriateness for the intended use of the ~~WDE,~~WQE, obvious errors, and individual outliers. Graphing of the data (true versus measured) is recommended as an assistive visual tool. ~~This graphic is available in the DQCALC software.~~

6.3.2 Outlying individual measurements must be evaluated; if determined to be erroneous, they should be eliminated using scientifically-based reasoning. Identification of potential outliers for data evaluation and validation may be accomplished using statistical ~~procedures, such as the optional one provided in the DQCALC software,~~ procedures or through visual examination of a graphical representation of the data. WQE computations must be based on retained data from at least six independent measurements at each of at least five concentration levels. The data removed and the percentage of data removed must be recorded and retained to document the WQE calculations.

6.4 *Modeling Standard Deviation ~~versus~~Versus True Concentration*—The purpose is to ~~characterize~~predict the ~~intralaboratory~~within-laboratory measurement standard deviation ~~(ILSD)~~(WLSD) as a function of true concentration, ~~σ = σ~~σ̂(T). ~~G (T).~~ The relationship is used for two purposes: (*1*) to provide weights (if needed) for fitting the mean-recovery model and (*2*) to provide the within-laboratory standard deviation estimates crucial to determining the WQEs.

NOTE 3—See Caulcutt and Boddy (**5**) for more discussion of standard deviation modeling and weighted least squares (WLS) in analytical chemistry.

6.4.1 This practice ~~utilizes~~uses four models as potential fits for the ~~IntraLaboratory~~Within-Laboratory Standard Deviation ~~(ILSD)~~(WLSD) model. The ~~identification~~selection process considers ~~(that is, fits and evaluates) each model in turn, from simplest to most complex, until a suitable model is found.~~a linear model first, performing statistical tests to decide whether a simpler constant model can be used or whether one of the more complicated curved models is required. See Carroll and Ruppert (**6**) for further discussion of standard-deviation modeling. The ~~model order is~~ models are as follows:

~~*Model A* (*Constant* ILSD *Model*): $s = g + \text{error}$~~ (1)

Model C (*Constant WLSD Model*): $s = g + \text{error}$ (1)

where:

~~g = a fitted constant.~~

s = standard deviation of measurement results, and
g = model parameter.

Under Model ~~A,~~C, standard deviation does not change with concentration, resulting in a relative standard deviation that declines with increasing concentration, T.

~~*Model B* (*Straight – line* ILSD *Model*): $s = g + h \times T + \text{error}$~~ (2)

Model L (*Linear WLSD Model*): $s = g + h \times T + \text{error}$ (2)

where:

~~g and h = fitted constants.~~

s = standard deviation of measurement results,
T = true concentration, and
g and h = model parameters.

Under Model ~~B,~~L, standard deviation increases linearly with concentration, resulting in an asymptotically constant relative standard deviation as T increases.

~~*Model C* (*Hybrid* ILSD *Model*): $s = \{g^2 + (h \times T)^2\}^{1/2} + \text{error}$~~ (3)

$$\textit{Model H (Hybrid WLSD Model)}: \quad s = \sqrt{g^2 + (h \times T)^2} + \text{error} \tag{3}$$

where:

~~*g* and *h* = fitted constants.~~

s = standard deviation of measurement results,
T = true concentration, and
g and *h* = model parameters.

Under Model ~~D,~~H, within-laboratory standard deviation increases with concentration in such a way that the relative standard deviation declines as *T* increases, approaching an asymptote of *h*.

~~$$\textit{Model D (Exponential ILSD Model)}: \quad s = g \times \exp(h \times T) + \text{error} \tag{4}$$~~

$$\textit{Model E (Exponential WLSD Model)}: \quad s = g \times \exp(h \times T) + \text{error} \tag{4}$$

where:

~~*g* and *h* = fitted constants.~~

s = standard deviation of measurement results,
T = true concentration, and
g and *h* = model parameters.

Under Model ~~D,~~E, within-laboratory standard deviation increases exponentially with concentration, resulting in a relative standard deviation that may initially decline as *T* increases, but eventually increases ~~as~~with *T*, ~~increases.~~so that there is at most a bounded quantitation range within which the RSD is less than *Z* / 100.

6.4.1.1 The procedures for estimating the parameters of each model and their uncertainties are given in Appendix X2 – Appendix X5. Model L should be tried first, since its fitting procedure (in Appendix X2) provides criteria for choosing among the various models.

6.4.1.2 In all cases, it is assumed that $g > 0$. A value of $g < 0$ has no practical interpretation and may indicate that a different ~~ILSD~~WLSD model should be used. Furthermore, it is assumed that *g* is not underestimated by censored data among measurements of blanks or other low-concentration samples. If $h < 0$, it must not be statistically significant, and Model ~~A~~C should be evaluated.

6.4.2 ~~The ASTM D19 Practice~~ If ~~D7510 describes the DQCALC software that can be used to perform the calculations for each of the four models, as well as the fit of each (this product can be obtained by contacting ASTM and asking for the DQCALC adjunct). The software identifies which model produced the best fit, and allows the user to select either this model or an alternative model. The software provides various graphical representations of the data and residuals, and the user manual provides assistance in using and interpreting the graphics and calculated values. Evaluation of the fit of each model to the data (as well as knowledge of chemistry, the method, and the systems used to generate the data) and judgment are important when selecting the most appropriate model. Where~~ a model other than the best fit is chosen, the reason for the choice should be scientific and should be recorded to document the WQE.

6.4.2.1 ~~Users of this practice not using the ASTM D19 DQCALC software can consult Practice D6091, which contains a protocol that provides the full procedural, consensus-balloted basis for these calculations. It is also recommended that those not using the software graph the relationship of true concentration to measurement standard deviation, and visually verify~~ It is recommended that the relationship of measurement standard deviation to true concentration be graphed and used for visual verification of the appropriateness of each model and of the model selected for use.

6.5 *Fitting the Mean-Recovery Relationship (Measured versus True Concentration)*—Based on the standard-deviation model selected (constant versus other models), the mean-recovery concentration is fitted versus true concentration, using ordinary least squares or weighted least squares, ~~respectively. The mean-recovery~~ as appropriate. The mean recovery is evaluated for statistical significance and lack of fit. A graph of mean recovery (along with the "calibration" line) and a graph of the residuals should also be visually examined. ~~The ASTM D19 DQCALC software performs these activities automatically. Alternatively, many~~ Many off-the-shelf statistical software packages may ~~also be used.~~ be used for this purpose.

NOTE 4—Regression coefficients should not be used to assess goodness of fit.

6.5.1 The mean-recovery regression (~~true~~(measured versus ~~measured~~true concentration) model is a ~~simple~~ straight line,

Model R: $Y = a + b\,T + \text{error}$ (5)

The fitting procedure depends on the ~~standard-deviation-model selection.~~ chosen standard-deviation model. If the constant model, Model ~~A,~~C, was selected, then ordinary least squares (OLS) can be used to fit Model R for mean ~~recovery (see the left column of~~ recovery. In all other cases, ~~Table 1, or Caulcutt and Boddy (5)). If a non-constant standard-deviation model was selected, then~~ weighted least squares (WLS) should be ~~used to fit mean recovery. The~~ used. WLS approximately provides the minimum-variance unbiased linear estimate of the ~~coefficients,~~parameters, *a* and *b*. The WLS procedure is described in ~~the~~Appendix X6 ~~IDE Practice D6091~~.

6.6 *Compute the WQE for each Z (%RSD)*—Using the mean-recovery regression line determined above, the most appropriate model of the relationship of relative standard deviation to true concentration (also determined above), and the *Z* value desired, the user obtains the WQE, which is ~~the~~ an estimate of the lowest true concentration (corresponding to the measured concentration) at which the desired ~~%~~ RSD ~~was achieved.~~is achieved. Procedures for calculating the WQE and estimating its relative standard uncertainty are given in Appendix X2 – Appendix X5.

6.6.1 Given the standard-deviation model, its estimated parameters, and the mean-recovery regression line, there is a lower limit, Z_{lim}, below which $WQE_{Z\,\%}$ cannot be calculated because there is no true concentration at which RSD equals *Z* %. For $Z > Z_{lim}$, the WQE can be calculated but its true value may nevertheless be extremely uncertain, especially when *Z* is near Z_{lim}. If the relative standard uncertainty, u_{rel}(WQE), exceeds 25 %, the calculated WQE should be considered an unreliable estimate. The actual reliability of the WQE also depends on the adequacy of the standard-deviation model, which is more difficult to quantify.

NOTE 5—Under Model C, Z_{lim} is zero. For the other models described here, Z_{lim} is always positive.

6.6.2 The measured concentration ~~(YQ)~~(Y_Q) at which the desired ~~% RSD~~ RSD was achieved may also be of interest for some uses. This value is the level at which the required ~~% RSD~~ RSD was obtained in measured concentration units (that is, the value, paired with a WQE, that has not been corrected for bias through the mean-recovery regression). Where the ~~*YQ*~~Y_Q and the WQE are equal (following application of significant figures and rounding), there is no apparent bias ~~present~~ at the WQE concentration.

~~6.6.2 The WQE is the lowest true concentration at which (based on the modeling of standard deviation at that concentration and including the required confidence for the sample size (90% tolerance interval)) the percent relative standard deviation is achieved at the desired Z. The DQCALC adjunct software calculates the 10 %, 20 %, and 30 % WQE as the typical Z values.~~

~~6.6.2.1 Fig. 1 provides an example that demonstrates a case with positive bias (intercept greater than zero) and imperfect recovery (slope of the calibration not equal to one), thereby highlighting the advantages of the WDE procedure. More simplistic quantitation procedures often make inappropriate assumptions about slope (that is, assume it to be one) and y-intercept (that is, assume it to be zero at a true concentration of zero), in addition to assuming that the standard deviation is constant. Additionally, where the simplest model (constant) for standard deviation is rejected, the WDE procedure requires that weighted least squares be used for fitting the recovery model, thus preventing higher concentrations from having an excessive effect on the resulting curve; most other practices do not offer this protection.~~

7. Review, Documentation and Reporting

7.1 The WQE analysis report should include: (*1*) the identification of laboratory and (*2*) identification of analytical method, analyte(s), matrix (or matrices), sample properties (for example, volume or mass) and specific method options (if any) utilized. Where the laboratory uses standard operating procedures (SOPs) to implement methods or method protocols, these SOPs should be referenced, including the identification of any revision/version. Documentation of each datum used should be equivalent to that of reported data (for example, instrument, analyst, date, etc.). There should be a description of all data-screening procedures employed, all results obtained, all individual values omitted from further analysis (that is, outliers that have been removed), all missing values, and the percentage of data utilized in the calculations relative to the initial data set. Any anomalies encountered should be listed, including and anomalous calibration or quality control sample results (for example, data validation qualifiers or flags). The data (statistical) analysis should be included or referenced ~~(for example, the output file from the DQCALC software)~~ and the WQE values determined recorded. The selected standard-deviation model, plus the coefficient estimates for this model and for mean-recovery model, should also be recorded. Where a statistical model other than the mathematical best fit has been chosen, the reasoning should be described.

8. Report

8.1 The analysis report should at a ~~miminum~~minimum contain:

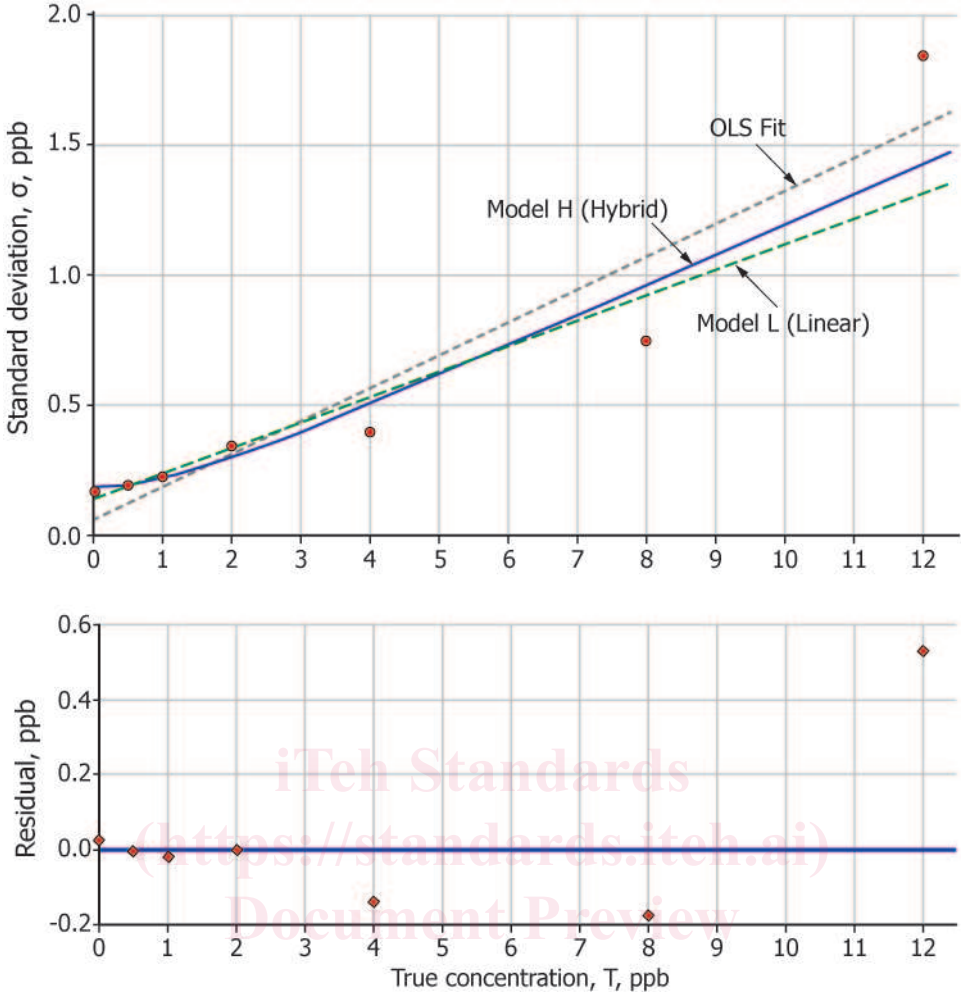

FIG. 1 Sample Standard Deviations ~~(+)~~ Versus True Concentration, with ~~Straight-Line~~Linear Fit, Hybrid Model Fit, and Residuals from ~~Straight-Line~~Linear Fit (Lower Plot), All in ppb

8.1.1 Identification of laboratory,

8.1.2 Analytical method,

8.1.3 Analyte(s),

8.1.4 Matrix (or matrices),

8.1.5 Sample properties (for example, volume),

8.1.6 Study design,

8.1.7 Analyst, method, and date of testing for each study sample,

8.1.8 Any anomalies in the study, including QA/QC sample results,

8.1.9 Data-screening results, individual values and laboratories omitted from further analysis, and missing values,

8.1.10 ~~IL SD~~WLSD model selected, and

8.1.11 Coefficient estimates for the ~~IL SD~~WLSD model and mean-recovery model.

~~NOTE 5—The DQCALC input and output files provide much of this documentation.~~

8.2 The report should be given a second-party review to verify that:

8.2.1 The data transcription and reporting have been performed correctly,

8.2.2 The analysis of the data and the application of this standard have been performed correctly, and

8.2.3 The results of the analysis have been used appropriately, including assessment of assumptions necessary to compute a WQE.

NOTE 6—Reviewer(s) should be qualified in one or ~~both~~more of the following areas: (*1*) applied statistics, ~~and~~ (*2*) metrology, and (*3*) analytical chemistry.

8.3 A statement of the review and the results of the review should accompany the report.

9. Rationale

9.1 The basic rationale for the WQE is contained in Currie (**1**). The WQE is a performance characteristic of an analytical ~~method,~~ measurement process, to paraphrase Currie. ~~As with the Within-Laboraotory Detection Estimate (WDE), the~~ Like an estimated detection limit, the WQE is helpful for the planning and use of chemical analyses. The WQE is another benchmark indicating whether the method can adequately meet measurement needs.

9.2 The idealized definition of $WQE_{Z\%\,\%}$ is that it is the lowest concentration, ~~LQ~~L_Q, that satisfies: ~~TL_Q~~ = ~~$(100/Z)$~~$(100/Z)\,\sigma_{L_Q}$ ~~$\leq_T$~~(where ~~$\leq\sigma_{L_Q T}$~~ is the actual standard deviation of ~~interlaboratory~~within-laboratory measurements at concentration ~~TL_Q~~); this definition is equivalent to ~~satisfying, % satisfying~~ RSD = $\sigma_{L_Q} \leq / L_Q$~~T = TZ = Z~~ %. In other words, $WQE_{Z\%\,\%}$ is the lowest concentration with ~~Z %~~Z % RSD (assuming such a concentration exists). If, as is commonly the case, ~~% RSD~~ RSD declines with increasing true concentration, then the relative uncertainty of any measurement of a true concentration greater than the WQE will not exceed ~~±Z %.~~±Z %. The range, ~~±3≤±3~~$_{LQ}\sigma_{L_Q}$, is an approximate prediction or confidence interval very likely to contain the measurement, which is assumed to be normally distributed. This assertion is based on critical values from the normal distribution (or from the student's *t* ~~distribution~~ -distribution if ~~≤σ~~ is estimated rather than known). Then, with high confidence, the relative error of any measurement of a true concentration greater than the WQE will not exceed ~~±3~~±3 × *Z* ~~· Z %.~~ %. For example, a measurement above the $WQE_{10\,\%}$ (and assumed to have true concentration above the WQE) could be reported as 6 ppb (±30 %) = ~~6 (±2)~~ (6 ± 2) ppb, with a high degree of certainty.

9.3 There are several real-world complications to this idealized situation. See Maddalone et al. (**7**), Gibbons (**8**), and Coleman et al. (**9**). Some of these complications are listed as follows:

9.3.1 Analyte recovery is not perfect; the relationship between measured values of concentrations and true concentrations cannot be assumed to be trivial. There is bias between true and measured values. Recovery can and should be modeled. Usually, a straight line will suffice. In practice, when both the standard deviation and recovery models are used, WQEZ % is calculated to be the lowest concentration L_Q that satisfies $L_Q = (100/Z)\hat{\sigma}(L_Q)/b$.

9.3.2 Variation is introduced by different laboratories, analysts, models, and pieces of equipment; environmental factors; flexibility/ambiguity in a test method; contamination; carryover; matrix influence; and other factors. It is intractable to model these factors individually, but their collective contributions to measurement ~~ILSD~~WLSD can be observed, if these contributions are part of how a study is designed and conducted.

9.3.3 The standard deviation of measurements is generally unknown, and may change with true concentration, possibly because of the physical principle of the test method. To ensure that a particular ~~%~~ RSD is attained at or above the WQE, there must be a way to predict the ~~ILSD~~WLSD at different true concentrations. Short of severely restricting the range of concentrations for a study, prediction is accomplished by an empirical ~~ILSD~~WLSD model. In all of the respects discussed in 9.1 – 9.3, $WQE_{10\,\%}$ is similar to the AML developed by Gibbons et al. (**10**). However, the AML follows an approximate approach, where the standard deviation used in the ~~10≤~~10σ formula is estimated at a detection critical value, and then is taken to be a constant (over a trace-level range of concentrations) for the ~~10≤~~10σ computation. In contrast, $WQE_{10\,\%}$ follows the "more statistically and conceptually rigorous" approach described by Gibbons et al. (**8**), and contained in Currie (**1**). This greater rigor comes at the risk of: (*a*) possibly being unattainable for some methods (for which only a less strict level of ~~%~~ RSD can be ensured); (*b*) having uncertainty that is potentially complex, and depends both on the model used and on the data.

10. Keywords

10.1 critical limits; matrix effects; precision; quantitation; quantitation limits