Designation: D 6512 – 00

Standard Practice for Interlaboratory Quantitation Estimate[1]

This standard is issued under the fixed designation D 6512; 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.

1. Scope

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

1.2 $IQE_{Z\,\%}$ is computed to be the lowest concentration for which a single measurement from a laboratory selected from the population of qualified laboratories represented in an interlaboratory study will have an estimated Z % relative standard deviation (Z % RSD, based on interlaboratory standard deviation), where Z is typically an integer multiple of 10, such as 10, 20, or 30, but Z can be less than 10. The $IQE_{10\,\%}$ is consistent with the quantitation approaches of Currie **(1)**[2] and Oppenheimer, et al **(2)**.

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

1.3.1 *Routinely Achievable IQE Value*—Most laboratories are able to attain the IQE quantitation performance in routine analyses, using a standard measurement system, at reasonable cost. This property is needed for a quantitation limit to be feasible in practical situations. Representative laboratories must be included in the data to calculate the IQE.

1.3.2 *Accounting for Routine Sources of Error*—The IQE should realistically include sources of bias and variation that are common to the measurement process. These sources include, but are not limited to: intrinsic instrument noise, some "typical" amount of carryover error; plus differences in laboratories, analysts, sample preparation, and instruments.

1.3.3 *Avoidable Sources of Error Excluded*—The IQE should realistically exclude avoidable sources of bias and variation; that is, those sources that can reasonably be avoided in routine field 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 a way to detect and either correct or eliminate them.

1.4 The IQE applies to measurement methods for which calibration error is minor relative to other sources, such as when the dominant source of variation is one of the following:

1.4.1 *Sample Preparation*, and calibration standards do not have to go through sample preparation.

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

1.4.3 *Differences in Laboratories* (for whatever reasons), perhaps difficult to identify and eliminate.

1.4.4 *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*—Typically, one would compute the lowest % RSD possible for any given dataset for a particular method. Thus, if possible, $IQE_{10\,\%}$ would be computed. If the data indicated that the method was too noisy, one might have to compute instead $IQE_{20\,\%}$, or possibly $IQE_{30\,\%}$. In any case, an IQE with a higher % RSD level (such as $IQE_{50\,\%}$) would not be considered, though an IQE with RSD <10 % (such as $IQE_{1\,\%}$) would be acceptable. The appropriate level of % RSD may depend on the intended use of the IQE.

2. Referenced Documents

2.1 *ASTM Standards:*

D 2777 Practice for Determination of Precision and Bias of Applicable Test Methods of Committee D-19 on Water[3]

D 6091 Practice for 99 %/95 % Interlaboratory Detection Estimate (IDE) for Analytical Methods with Negligible Calibration Error[3]

E 1763 Guide for Interpretation and Use of Results from

[1] This practice is under the jurisdiction of ASTM Committee D19 onWater and is the direct responsibility of Subcommittee D19.02 on General Specifications, Technical Resources, and Statistical Methods.

Current edition approved Feb. 10, 2000. Published May 2000.

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

[3] *Annual Book of ASTM Standards,* Vol 11.01.

Interlaboratory Testing of Chemical Analysis Methods[4]

3. Terminology

3.1 *Z % Interlaboratory Quantitation Estimate ($IQE_{Z\ \%}$)*, also denoted "LQ," for "Limit of Quantitation" in accordance with Currie **(1)**—The lowest concentration for which a single measurement from a laboratory selected from the population of qualified laboratories represented in an interlaboratory study will have an estimated Z % relative standard deviation (Z % RSD, based on interlaboratory standard deviation).

3.2 *Definitions of Terms Specific to This Standard:*

3.2.1 *Censored Measurement*—A measurement that is not reported numerically nor is reported missing, but is stated as a nondetect or a less-than (for example, "less than 0.1 ppb"). There are two reasons why the measurement may not be reported numerically. Either the measurement was considered insufficiently precise or accurate (these kinds of data should not be censored), or the identification of the analyte was suspect (these kinds of data should be censored). See §6.2.3.1. A reported "less than" may have the same meaning as a non-reported measurement, but a reported "less than" also implies (perhaps erroneously) that any concentration greater than or equal to the accompanying value (for example, 0.1 ppb) can be measured, and will be reported numerically.

3.2.2 *Quantitation Limit (QL) or Limit of Quantitation (LQ)*—A numerical value, expressed in physical units or proportion, intended to represent the lowest level of reliable quantitation. The IQE is an example of a QL.

4. Summary of Practice

4.1 Every ASTM Committee D-19 test method is evaluated to determine precision and bias by conducting a collaborative study, in accordance with Practice D 2777. That study, or a similar collaborative study, can also be used to evaluate the lowest concentration level of reliable quantitation for a test method, referred to herein as the interlaboratory quantitation estimate (IQE). Such a study must include concentrations suitable for modeling the uncertainty of mean recovery of interlaboratory measurement, preferably without extrapolation. The study must also be planned and conducted to allow the known, routine sources of measurement variability to be observed at typical levels of influence. After the study is conducted, outlying laboratories and individual measurements should be eliminated, using an accepted, scientifically based procedure for outlier removal, such as found in Practice D 2777. The IQE computations must be based on retained data from at least six independent laboratories at each concentration level.

4.2 Retained data are analyzed to identify and fit one of three proposed interlaboratory standard deviation (ILSD) models. These models describe the relationship between the interlaboratory standard deviation of measurements and the true concentration, *T*. The identification process involves evaluating the models in order, from simplest to most complex: constant, straight-line, and hybrid (proposed by Rocke and Lorenzato **(3)**). Evaluation includes statistical significance and residual analysis.

4.3 The chosen model is used to predict the standard deviation of interlaboratory measurements at any true concentration within the study concentration range. If interlaboratory standard deviations change systematically with respect to the true concentration (that is, they are NOT constant), the predictions are used to generate weights for fitting the mean-recovery relationship (the assumed straight-line relationship between measured concentration and true concentration), using weighted least squares. (Otherwise, ordinary least squares is used.) The mean-recovery curve is evaluated for statistical significance, for lack of fit, and for residual patterns. The ILSD model is also used to estimate the interlaboratory standard deviation at concentrations within the concentration range. Either a direct or interactive algorithm (depending on the model) is used to compute $IQE_{10\ \%}$, the lowest concentration with estimated RSD = 10 % (Z = 10). If there is no such concentration, then $IQE_{20\ \%}$ is computed instead, or $IQE_{30\ \%}$, if necessary. If supported by the data quality objectives (DQOs), $IQE_{Z\ \%}$ may be computed for some Z < 10.

5. Significance and Use

5.1 Appropriate application of this practice should result in an IQE achievable by most laboratories properly using the test method studied. That is, most laboratories should be capable of measuring concentrations greater than $IQE_{Z\ \%}$ with RSD = Z % or less. The IQE provides the basis for any prospective use of the test method by qualified laboratories for reliable quantitation of low-level concentrations of the same analyte as the one studied in this practice, and same media (matrix).

5.2 The IQE values may be used to compare the quantitation capability of different methods for analysis of the same analyte in the same matrix. The IQE is not an indicator of individual laboratory performance.

5.3 The IQE procedure should be used to establish the interlaboratory quantitation capability for any application of a method where interlaboratory quantitation is important to data use. The intent of the IQE is not to set reporting limits.

6. Procedure

6.1 The following procedure has stages described in the following paragraphs: 6.2–IQE Study Plan, Design, and Protocol; 6.3–Conduct the IQE Study, Screen the Data, and Choose a Model; and 6.4–Compute the IQE. A flowchart of the procedure is shown in Fig. 1.

6.2 *IQE Study Plan, Design, and Protocol*:

6.2.1 *Choose Analyte, Matrix, and Method*—At least one analyte of interest is selected, typically one for which there is interest in trace or near-trace levels of concentration, such as toxic materials that are controlled and regulated. For each analyte, an approximate maximum true concentration is selected, based on these considerations:

6.2.1.1 The anticipated IQE should be exceeded by a factor of 2 or more,

6.2.1.2 A single model, (ideally a straight-line model in true concentration, *T*) should describe mean recovery (that is, mean

[4] *Annual Book of ASTM Standards* , Vol 03.06.

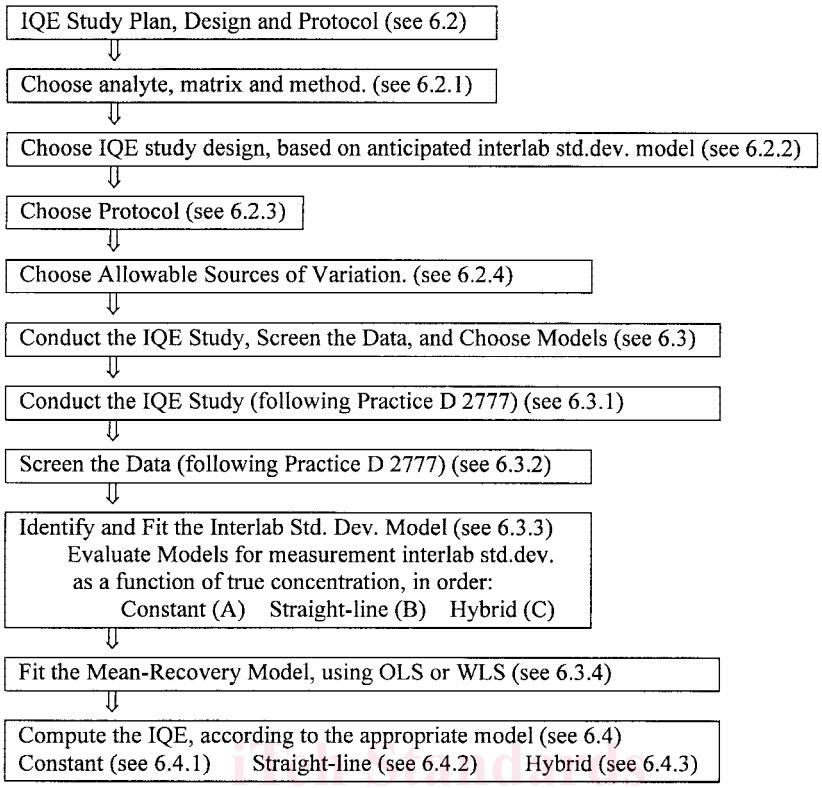

FIG. 1 Flowchart of IQE Procedure

measured concentration) for the entire range of concentrations, from zero to the selected maximum concentration.

NOTE 1—The IQE procedure uses the straight-line model for mean recovery, thus implicitly assuming that a straight line is adequate. Thus, the IQE would not be appropriate for cases where this assumption is unreasonable. For example, it would not hold for cases where there was systematic bias for most or all laboratories, such as a tendency to report values that are too high for some portion of the concentration range.

6.2.1.3 A single model in true concentration should describe the standard deviation of interlaboratory measurements for the entire range of concentrations, from zero to the selected maximum concentration.

6.2.1.4 The concentration range must be sufficient to enable statistically significant coefficients to be estimated for the ILSD model and mean-recovery model. At least one matrix of interest is also selected, and an accepted standard analytical method for those analytes is selected for study. If there is no possibility of matrix interference, then it may only be necessary to determine a list of acceptable matrices that can be used, instead of selecting a specific matrix. For example, for a particular analyte, concentration range, and method, it may be supposed that reagent waters from different laboratories are indistinguishable. However, that assumption may not hold for another analyte or another concentration range.

6.2.2 *Choose IQE Study Design*—The design should be based (if possible) on an anticipated ILSD model. Section 7 of Practice D 2777 can be followed for the study design and protocol. The anticipated form of the ILSD model (the relationship between interlaboratory measurement standard deviation and true concentration) can help in choosing an IQE study design. Three models are proposed herein for the relationship between the interlaboratory standard deviation of measurements and the true concentration: constant, straight-line (increasing), and hybrid (increasing). See 6.3.3 for details. Chemistry, physics, empirical evidence, or informed judgment may make one model more plausible than others. However, it may not be possible to anticipate the relationship between standard deviation and true concentration.

6.2.2.1 Select an IQE study design that has enough distinct concentration levels to assess statistical lack of fit of the models (see Draper and Smith **(4)**). Recommended designs are: (*a*) the "semi-geometric" design at five or more true concentrations, {T_1, T_2, and so forth}, such as: {0, $IQE_0/4$, $IQE_0/2$, IQE_0, $2 \times IQE_0$, $4 \times IQE_0$, $8 \times IQE_0$}, where IQE_0 is an initial estimate of the IQE (such as $10s'$ where s' is the interlaboratory measurement standard deviation at a trace-level, nonzero concentration); (*b*) equi-spaced design: {0, $IQE_0/2$, IQE_0, $(3/2) \times IQE_0$, $2 \times IQE_0$, $(5/2) \times IQE_0$, $3 \times IQE_0$}; and (*c*) any other design with at least five concentrations, provided that the design includes at least one concentration approximately equal to $2 \times IQE_0$, at least one nonzero concentration below IQE_0, and one blank, or unspiked sample. Preferably, the design will have at least seven concentrations, including a blank.

6.2.2.2 The study's concentration levels must either be known (true 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 (for example, the median value from many laboratories, or results from a recognized laboratory, such as NIST, using a high-accuracy method).

6.2.3 *Choose Protocol*—The protocol should follow Section 7 of Practice D 2777. The protocol should include design run order and details on when the system is to be purged, have extra blanks run, and so on. It should take into consideration possible problems with carryover, study cost (in time and money), and the time constants of drift of the measurement system or degradation of the sample.

6.2.3.1 For purposes of the collaborative study, the study supervisor should provide instructions to participating laboratories to disable (if possible) any internal reporting limits or any other data-censoring thresholds (such as an "instrument detection limit") that are used to determine whether a numerical measurement is to be reported as a number, or as a nondetect or less-than (that is, the number is censored). If censoring is unavoidable, the laboratory censoring threshold must be reported with the study data. However, qualitative criteria used by the method to identify and discriminate among analytes are separate criteria, and must be satisfied in accordance with the method.

6.2.4 *Choose Allowable Sources of Variation*—It is assumed that, collectively, the many sources of variation will cause interlaboratory measurements at any true concentration to be Normally distributed. The number of laboratories providing usable data must be maximized in order for the study to capture representative between-laboratory variation adequately. Ordinary within-laboratory variation must be allowed to affect the measurement process, as happens in routine measurement. Ideally, there would be many laboratories, and each measurement at each laboratory would be made as a routine measurement, made by a different analyst using a different (qualified) measurement system on a different day, in random order, without the analyst being aware of the true value, or even that the sample was part of a special study.

6.2.4.1 As emphasized in Practice D 2777, maximizing the number of participating laboratories is often the most important thing that can be done to guarantee a successful study. The number of laboratories providing a full set of usable data will typically fall short of the number of participating laboratories. A minimum of ten participating laboratories is recommended.

6.2.4.2 To the extent possible, the study should be conducted so as to mimic routine laboratory measurement, particularly if the method is labor-intensive, as opposed to a highly automated method. That is, not only should the analysts not be aware of the true concentrations of these samples, but also they should not know that they are measuring special, study samples. These restrictions minimize the risk of extra-care distortion of data so common in analytical studies. However, it is recommended that the participating analysts be told to disable data-censoring limits, because there may or may not be some low concentrations in the study samples (see 6.2.3.1).

6.2.4.3 For each laboratory, the maximum possible number of qualified analysts should be involved in the study, since there are variations that may be allowed by the method, may be practiced by different analysts, and will be seen in routine analyses.

6.2.4.4 For each laboratory, the maximum possible number of qualified measurement systems should be used, since there are model-to-model and instrument-to-instrument differences in equipment and maintenance, as will be seen in routine analyses.

6.2.4.5 For each laboratory, the IQE study should be scheduled to span the maximum possible number of days consistent with holding-time constraints, since day-to-day changes in analytical laboratory environmental conditions, contamination, solvent purity, and other factors can affect measurements, and will be seen in routine analyses.

6.3 *Conduct the IQE Study, Screen the Data, and Choose a Model*:

6.3.1 The IQE study should be conducted in accordance with Section 9 of Practice D 2777. Blank correction should not be performed by the laboratories, unless the method requires this subtraction in order to perform the test. Each laboratory should supply method-blank data along with the uncorrected measurement values, and the study supervisor can determine whether the reported measurements should be corrected.

6.3.2 The IQE study data should be screened in accordance with the initial subsections (relating to removing data) of Section 10 of Practice D 2777. (Proceed to Section 6.5 of the IQE Practice if, for any concentration, more than 10 % of the retained measurements are nondetects or less-thans.)

6.3.3 *Identify and Fit the ILSD Model*—The ILSD model should be identified and its coefficients should be estimated by using the following procedure. See Draper and Smith **(4)** and Caulcutt and Boddy **(5)** for more discussion of how to model standard deviations and how to do weighted least squares (WLS) in analytical chemistry. See Carroll and Ruppert **(6)** for further discussion of standard-deviation modeling. The ILSD model is an attempt to characterize the unknown (or partly known) relationship ($\sigma = G(T)$) between the actual standard deviation of interlaboratory measurement and true concentration. The model is used for two purposes: to provide weights for the WLS regression to fit the mean-recovery model, and to provide the interlaboratory standard-deviation estimates required to determine the IQE.

6.3.3.1 Three ILSD models are proposed. The identification process considers (that is, fits then evaluates) each model in turn, from simplest to most complex, until a suitable model is found. Prior knowledge can be combined with empirical results to influence the selection of a model if a suitable referenced publication can be cited. The model order is as follows:

(*a*) *Constant Model for the ILSD (Model A):*

$$s = g + \text{error} \quad (1)$$

where:

s = the sample standard deviation for interlaboratory measurements,
g = estimated constant, and

"error" is included for arithmetic completeness, since the model will not hold exactly. Interlaboratory standard deviation does not change with concentration, resulting in a relative standard deviation that declines with increasing T

(*b*) *Straight-line Model for the ILSD (Model B):*

$$s = g + hT + \text{error} \quad (2)$$

where: g and h = fitted constants.

Interlaboratory standard deviation increases linearly with concentration, resulting in an asymptotically constant relative standard deviation as T increases.

(*c*) *Hybrid Model for the ILSD (Model C):*

$$s = (g^2 + [hT]^2)^{(1/2)} + \text{error} \quad (3)$$

where the positive square root is taken; g and h are fitted constants. Interlaboratory standard deviation increases with concentration, at first slowly, then achieving proportional increase. This behavior also results in a relative standard deviation that initially declines as the concentration increases from zero, then asymptotically approaches a constant level. The Hybrid Model, the form of which was developed by Rocke and Lorenzato **(3)** is so-named because it incorporates two things: additive error with constant standard deviation (coefficient g), and multiplicative error with increasing standard deviation (coefficient h).

NOTE 2—The Hybrid Model used the form of Roche and Lorenzato, but not necessarily the same assumptions for error distribution. The Hybrid Model is also the same as the General Analytical Error Model of Guide E 1763

In all cases, it is assumed that $g > 0$ (though this constraint is irrelevant for the Hybrid Model). A value of $g < 0$ has no practical interpretation, and may indicate that a different ILSD model should be used. Furthermore, it is assumed that g is not underestimated because of censored data among measurements of blanks or other low-concentration samples. (Censoring is addressed in 6.2.3.1, 6.3.2, and 6.5).

If $h < 0$, then it must be significantly less than zero (statistically), in which case the Constant Model (Model A) should be evaluated.

6.3.3.2 *ILSD-Model Identification and Fitting Procedure*:

See Section 10 for a detailed example, using the Hybrid Model for the ILSD.

(*a*) Merge all retained IQE study data (after possible elimination of some data in accordance with 6.3.2).

(*b*) For each true concentration, T, compute the adjusted interlaboratory sample standard deviation, s_k, an estimate of the true underlying interlaboratory measurement standard deviation, σ_k. The adjusted interlaboratory sample standard deviation is the sample standard deviation, s_k , multiplied by the bias-correction factor, a'_n, found in Table 1. In this Practice, all references to computed and fitted values of the interlaboratory sample standard deviation refer to adjusted values.

(*c*) Plot s_k versus T_k.

(*d*) Using ordinary least squares (OLS, see Caulcutt and Boddy **(5)**), regress s_k on T_k, temporarily assuming that the Straight-line Model is valid. The regression provides coefficients, g and h, in the relationship,

TABLE 1 Bias-Correction Adjustment Factors for Sample Standard Deviations Based on n Measurements (at a particular concentration)[A]

n	2	3	4	5	6	7	8	9	10
a'_n	1.253	1.128	1.085	1.064	1.051	1.042	1.036	1.031	1.028

[A]For each true concentration, T_k, the adjusted value $s_k = a'_n s'_k$ should be modeled in place of sample standard deviation, s'_k. For $n > 10$, use the formula, $a'_n = 1 + [4(n-1)]^{-1}$. See Johnson and Kotz **(7)**.

$$s_k = g + hT_k + \text{error}. \quad (4)$$

Compute residuals,

$$r_k = s_k - (g + h\,T_k) \quad (5)$$

Plot r_k versus T_k.

(*e*) Evaluate the reasonableness of the Constant Model for the ILSD (Model A) as follows: First, note the p-value associated with slope estimate, h, from the OLS regression. If the p-value is less than 5 %, there is statistically significant slope, and the Constant model should be rejected; proceed to the next step. Second, examine the plots produced in (*c*) and (*d*). If obvious systematic curvature is present (for example, quadratic-like behavior), both the Constant Model and the Straight-line Model should be rejected; proceed to (*i*). If the Constant Model is not rejected, proceed to 6.3.4.

(*f*) The Constant Model (Model A), has been rejected because of statistically significant slope. Evaluate the reasonableness of the Straight-line model for the ILSD (Model B). Examine the plot produced in (*d*). If obvious systematic curvature is present (for example, quadratic-like behavior), with a minimum that appears to be in the concentration range, the Straight-line Model should be rejected; proceed to (*j*). If the Straight-line Model is not rejected by this examination, proceed to 6.3.4, or, optionally, conduct a formal test for curvature, as follows in (*g*) through (*i*) (note that the usual and more general lack-of-fit test is not applicable for this modeling effort because there are no replicate sample standard deviations, s_k, for any concentration).

(*g*) Using OLS, regress T_k^2 on T_k, producing fitted coefficients u and v, used only to compute residuals, q_k, which comprise the orthogonal component of the quadratic term, T_k^2:

$$q_k = (\textit{predicted}\ T_k^2) - T_k^2 = (u + vT_k) - T_k^2 \quad (6)$$

(*h*) Using OLS, regress s_k on T_k and q_k simultaneously, producing fitted coefficients g and h (as before), but additionally Q:

$$s_k = g + hT_k + Qq_k + \text{error} \quad (7)$$

The only results of interest are the statistical significance and the sign of Q. These results collectively indicate the strength of evidence for curvature.

(*i*) Note the p-value, p_Q, associated with Q. Because q_k is orthogonal to T_k, this p-value indicates the level of statistical significance of (quadratic) curvature.

NOTE 3—Even though the test for curvature uses a quadratic term, a quadratic model is not one of the three recommended model choices. If $p_Q < 5$ % and $Q > 0$, there is sufficient statistical evidence of curvature in the relationship between s_k and T_k to warrant the use of the Hybrid Model, Model C ($Q > 0$ ensures that the increase in s_k with respect to T_k is faster than linear). If these conditions do not hold, then the Straight-line Model (Model B) is the appropriate model to use. Proceed to 6.3.4

(*j*) The Hybrid Model for the ILSD (Model C) can be used if there is evidence of curvature.

(*k*) To evaluate the reasonableness of the Hybrid Model, Model C, the model must first be fitted using nonlinear least squares (NLLS), either by Newton's-Method iteration (presented in the appendix), or another NLLS method.

(*l*) The fit from the Hybrid Model should be evaluated. A plot of the residuals, in log form, should be constructed: plot r_k versus T_k, where:

$$r_k = \ln s_k - \ln \hat{s}_k, \quad (8)$$

and $\hat{s}_k$ is the predicted value of s_k using the model. The plot should show no systematic behavior (for example, curvature). If the fit satisfies both types of evaluation, go to 6.3.4. Otherwise, a different (and possibly more complex) model may be used, such as the exponential model: $s = g \exp\{hT\}\cdot(1 +$ error). If there are enough true concentrations, a model with more coefficients could be considered; possibilities include quadratic (strictly increasing with increasing concentration), or even cubic.

6.3.4 *Fit the Mean-Recovery Model*—The mean-recovery model is a simple straight line,

$$\text{Model R: } Y = a + b\,T + \text{error}. \quad (9)$$

The fitting procedure depends on the model selection from 6.3.3. If the constant model, Model A, was selected for ILSD, then OLS can be used to fit Model R for mean recovery (see the left column of Table 2, or Caulcutt and Boddy **(5)**). If a nonconstant ILSD model was selected, such as the Straight-line Model (Model B), or the Hybrid Model (Model C), then weighted least squares (WLS) should be used to fit mean recovery. The WLS approximately provides the minimum-variance unbiased linear estimate of the coefficients, *a* and *b*. The WLS procedure is described in 6.3.4.1

6.3.4.1 *Weighted Least Squares Procedure, Using the Interlaboratory Standard Deviation (ILSD) Model*:

(*a*) Using the ILSD model and coefficient estimates from 6.3.3, compute the predicted interlaboratory standard deviation, $\hat{s}_k$, for each true concentration, T_k:

$$\text{Model B: } \hat{s}_k = g + h\,T_k \quad (10)$$

$$\text{Model C: } \hat{s}_k = (g^2 + [hT_k]^2)^{(1/2)} \quad (11)$$

(*b*) Compute weights for WLS:

$$w_k = (\hat{s}_k)^{-2}. \quad (12)$$

Note that if WLS is carried out using computer software, the default setting for weights may be different. For example, instead of supplying the values,
$(\hat{s}_k)^{-2}$, as weights, the software may require the user to supply values $(\hat{s}_k)$ or $(\hat{s}_k)^2$ as weights that are internally transformed by the software.

TABLE 2 Ordinary Least Squares (OLS) and Weighted Least Squares (WLS) Computations to Estimate Straight-line Model Coefficients
(Computations shown for convenience and contrast)

OLS	WLS
$\bar{T} = \frac{1}{n}\sum_{i=1}^{n} T_i$,	$\bar{T}_w = \frac{1}{n}\sum_{i=1}^{n} w_i T_i$,
$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$	$\bar{y}_w = \frac{1}{n}\sum_{i=1}^{n} w_i y_i$
$S_{TT} = \sum_{i=1}^{n} (T_i - \bar{T})^2$	$S_{wTT} = \sum_{i=1}^{n} w_i (T_i - \bar{T})^2$
$S_{TY} = \sum_{i=1}^{n} (T_i - \bar{T})(y_i - \bar{y})$	$S_{wTY} = \sum_{i=1}^{n} w_i (T_i - \bar{T})(y_i - \bar{y})$
slope = $b = S_{TY}/S_{TT}$	slope = $b = S_{wTY}/S_{wTT}$
intercept = $a = \bar{y} - b\bar{T}$	intercept = $a = \bar{y}_w - b\,\bar{T}_w$

(*c*) Carry out WLS computations analogous to OLS computations. See Table 2 or Caulcutt and Boddy **(5)**. The result will be coefficient estimates, *a* and *b*, for the mean-recovery model, Model R. Appendix II describes three approximate approaches to WLS commonly practiced, but not acceptable for this application.

(*d*) After fitting, the mean-recovery model should be evaluated for reasonableness and lack of fit. This evaluation should be done by ensuring the following: (*1*) The fit is statistically significant (overall *p*-value < 5 %); (*2*) The lack-of-fit *p*-value (if available; see Caulcutt and Boddy **(5)** or Draper and Smith **(4)**) is not statistically significant (lack-of-fit *p*-value >5 %); (*3*) A plot of the residuals shows no obvious systematic curvature (for example, quadratic-like behavior). If the mean-recovery model fails the evaluation, then the study supervisor will have to determine if only a subset of the data should be analyzed (perhaps the model fails for the higher concentration(s)), or if more data are needed.

6.4 *Compute the IQE*—The IQE is computed using the ILSD model to estimate the interlaboratory standard deviation, and using the mean-recovery model to scale the standard deviation. For any computed IQE to be valid, it must lie within the range of concentrations used in the study. The general form of the computation is to find the solution, LQ (within the range of concentrations used in the study), to the following equation:

$$T = (100/Z)\cdot G(T) \quad (13)$$

where function *G(T)* is the estimated interlaboratory standard deviation (in concentration units) of true value, T, and *Z* is taken to be 10, 20, or 30, in increasing order. That is, the first attempt is to compute $IQE_{10\,\%}$. If $IQE_{10\,\%}$ does not exist or is outside the range of concentrations used in the study, then $IQE_{20\,\%}$ is computed, if possible. If $IQE_{20\,\%}$ does not exist or is outside the range of concentrations used in the study, then $IQE_{30\,\%}$ is computed, if possible. If appropriate for a particular use, $IQE_{Z\,\%}$ can be computed for any value of Z <10, but Z>30 is not recommended. Thus, the IQE computation depends on the form of the ILSD model, which is part of function G. The ratio, $Z' = 100\cdot h/b$, represents the limit of the %RSD achievable. Therefore the strictest IQE achievable by the analytical method studied is $IQE_{Z'\,\%}$. For example, if $Z' = 100\cdot 0.17/1.0 = 17$, then the strictest IQE achievable would be the $IQE_{20\,\%}$(according to the nearest higher multiple of 10).

6.4.1 *ILSD Constant Model (Model A)*—In this case, $\hat{s} = g$; hence $G(T) = g/b$ and $LQ = (100/Z)\cdot g/b$. Thus,

$$IQE_{Z\,\%} = (100/Z)\cdot g/b \quad (14)$$

6.4.2 *ILSD "Straight-line" Model (Model B)*—In this case, $\hat{s} = g + hT$; hence $G(T) = (g + h\,T)/b$. To find the IQE, one must solve for *T*: $T = (100/Z)\cdot(g + h\,T)/b$. The solution is:

$$IQE_{Z\,\%} = g/(b\cdot(Z/100) - h). \quad (15)$$

6.4.3 *ILSD Hybrid Model (Model C)*—(additive and multiplicative error, in accordance with Rocke and Lorenzato **(3)**). In this case, $\hat{s} = (g^2 + [h\cdot T]^2)^{(1/2)}$; hence $G(T) = (g^2 + [h\cdot T]^2)^{(1/2)}/b$. To find the IQE, one must solve

$$T = ((100/Z)/b)\,(g^2 + [h\cdot T]^2)^{(1/2)} \quad (16)$$