Designation: E1763 – 06

Standard Guide for
Interpretation and Use of Results from Interlaboratory Testing of Chemical Analysis Methods[1]

This standard is issued under the fixed designation E1763; 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 guide covers procedures to help a task group interpret interlaboratory study (ILS) statistics to state precision and accuracy of a test method and make judgments concerning its range of use.

1.2 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

2. Referenced Documents

2.1 *ASTM Standards:*[2]

D6512 Practice for Interlaboratory Quantitation Estimate

E135 Terminology Relating to Analytical Chemistry for Metals, Ores, and Related Materials

E1601 Practice for Conducting an Interlaboratory Study to Evaluate the Performance of an Analytical Method

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

E1914 Practice for Use of Terms Relating to the Development and Evaluation of Methods for Chemical Analysis

3. Terminology

3.1 For definitions of terms used in this guide, refer to Terminology E135.

3.2 For descriptions of terms used in this guide, refer to Practice E1914.

4. Significance and Use

4.1 A written test method is subjected to an ILS to evaluate its performance. The ILS produces a set of statistical estimates that depend upon the method, but also are influenced by the laboratories and test materials involved in the study. For that reason, the ILS task group must interpret these estimates, aided by this guide and using analytical judgment, to decide if the method is suitable to be balloted for publication as a standard. The task group may use this guide to help them prepare the precision and bias statements that are a required part of the method.

5. Interlaboratory Studies

5.1 The following statement is required in each test method:

5.1.1 This test method has been evaluated in accordance with Practice E1601 and Guide E1763. Unless otherwise noted in the precision and bias section, the lower limit in the scope of each method specifies the lowest analyte content that may be analyzed with acceptable error (defined as a nominal 5 % risk of obtaining a 50 % or larger relative difference in results on the same test sample in two laboratories).

NOTE 1—this guide has been found to be useful for calculation of the interlaboratory quantitation estimate in accordance with Practice D6512, using known reference values in place of the measured values for application to that practice. It has been found necessary to replace the zero value for the blank with a very small number in order to apply this practice to that method to avoid the divide by zero computational error.

6. Required Statistical Information

6.1 A task group satisfies the requirement for statistical information if the method includes a table of the ILS statistics prepared in accordance with 6.2 and 6.3 and a summary statement selected from the model statements in Section 9. If the task group wishes to provide further statistical information, it may do so in accordance with the provisions of Section 7.

6.2 *Variability Data*—List the variability statistics for each analyte in a separate table arranged by increasing analyte content. List the number of independent data sets used in the calculations and the ILS statistics calculated in accordance with Practice E1601. Where appropriate, list the material type and reference material identification. Follow the examples of Table 1, Table 2, and Table 3.

---

[1] This guide is under the jurisdiction of ASTM Committee E01 on Analytical Chemistry for Metals, Ores and Related Materials and is the direct responsibility of Subcommittee E01.22 on Laboratory Quality.

Current edition approved Nov. 1, 2006. Published November 2006. Originally approved in 1995. Last previous edition approved in 2003 as E1763 – 98 (2003). DOI: 10.1520/E1763-06.

[2] 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.

TABLE 1 Gold in Bullion by the Fire Assay Method

Test Material	Number of Labor-atories	Gold found, %	Minimum SD (s_M, E1601)	Repro-ducibility SD (s_R, E1601)	Repro-ducibility Index (R, E1601)	$R_{rel\%}$
6	7	26.350	0.0089	0.0318	0.089	0.34
4	10	65.744	0.0236	0.0439	0.123	0.25
2	10	73.831	0.0261	0.0296	0.083	0.20
3	10	76.484	0.0275	0.0543	0.152	0.19
1	10	78.392	0.0200	0.0689	0.193	0.11
5	7	99.060	0.0189	0.0368	0.103	0.10

TABLE 2 Manganese in Iron Ores by the Permanganate Titrimetric method

Test Material	Number of Labora-tories	Man ganese found, %	Minimum SD (s_M, E1601)	Repro-ducibility SD (s_R, E1601)	Repro-ducibility Index (R, E1601)	$R_{rel\%}$
1	8	0.62	0.0047	0.0069	0.0193	3.11
2	8	1.17	0.0189	0.0219	0.0614	5.25
3	8	1.72	0.0237	0.0244	0.0683	3.97
4	8	2.83	0.0218	0.0244	0.0683	2.41
5	8	3.73	0.0218	0.0360	0.1007	2.70
6	8	5.55	0.0275	0.0724	0.2026	3.65

TABLE 3 Boron in Steel by the Curcumin Spectrophotometric Method

Test Material	Number of Labora-tories	Boron found, %	Minimum SD (s_M, E1601)	Repro-ducibility SD (s_R, E1601)	Repro-ducibility Index (R, E1601)	$R_{rel\%}$
1–D1,1	14	0.00023	0.000036	0.000064	0.00018	78.3
2–B1,2	21	0.00023	0.000082	0.000102	0.00028	124
3–B1,1	21	0.00026	0.000046	0.000084	0.00024	90.4
4–D1,7	14	0.00045	0.000046	0.000150	0.00042	63.3
5–B1,3	21	0.00046	0.000061	0.000107	0.00030	65.2
6–D1,2	14	0.00108	0.000054	0.000100	0.00028	25.9
7–B1,4	21	0.00136	0.000068	0.000189	0.00053	39.0
8–D1,3	14	0.00275	0.000104	0.000129	0.00036	12.1
9–D1,4	14	0.00315	0.000104	0.000129	0.00036	11.4
10–B1,5	21	0.00362	0.000111	0.000214	0.00060	13.3
11–D1,5	14	0.00378	0.000104	0.000257	0.00072	19.0
12–B1,6	21	0.00432	0.000143	0.000189	0.00053	12.3
13–D1,8	14	0.00432	0.000096	0.000171	0.00048	11.1
14–D1,6	14	0.00639	0.000132	0.000471	0.00132	15.2
15–B1,7	21	0.00904	0.000179	0.000482	0.00135	14.9
16–B1,8	21	0.0114	0.00035	0.000625	0.00175	15.4

	Certified Boron, %	B-value, %	Material Identification (Source)	Description
1	...	...	...	non-alloyed steel
2	0.0003	–0.00007	ERMC 097-1	high purity iron
3	0.0003	–0.00007	ERMC 283-1	high speed steel
4	...	...	...	alloyed steel
5	0.0004	0.00006	BAM 187-1	low alloyed steel
6	...	...	...	non-alloyed steel
7	0.0015	–0.00014	BCS 456/1	mild steel
8	...	...	...	non-alloyed steel
9	...	...	...	non-alloyed steel
10	0.0038	–0.00018	BAM 284-1	stainless steel
11	...	...	...	non-alloyed steel
12	0.0041	0.00022	BAM 178-1	low alloyed steel
13	...	...	...	alloyed steel
14	...	...	...	non-alloyed steel
15	0.0090	–0.00004	JSS 175-5	mild steel
16	0.0118	–0.0004	BCS 459/1	carbon steel

6.2.1 If the analytical method includes optional conditions extending the analyte range (for example, decreased sample portions or multiple calibration curves,) list the ILS statistics for each option separately as shown in Table A3.1 and Table A3.3.

6.3 *Bias Data*—If the ILS includes one or more test materials having an accepted reference value, include the accepted value(s) and b-value(s) as shown in Table 3.

7. Models for Error in Analytical Methods

7.1 An estimate of the reproducibility index, *R*, is obtained in an ILS for each individual test material. These are the discrete values of *R* listed in a statistical information table. Users need an estimate of *R* at the analyte level, which may lie anywhere within the range of the scope of the test method. If the task group conducted the ILS properly and employed good quality test materials having compositions that cover the application range, that information may be provided by following the procedures in this section.

7.1.1 If the analytical method includes sample portion or calibration options, treat the ILS statistics for each option as a separate method.

7.2 The task group must decide if the statistics for the test materials in the ILS exhibit trends that follow one of the three error models included in this section. The use of a model is essential if the task group intends to describe the behavior of the reproducibility index, *R*, as a function of analyte content. If the task group cannot agree on one model, it should not attempt to relate *R* to analyte content. The keys to identifying the model are the trends in *R* and $R_{rel\%}$ as the analyte content increases. Annex A1 includes a more detailed discussion of these analytical error models.

7.3 *General Model for Error in Analytical Methods*—The ILS data follows the general model if, with increasing analyte concentration, *R* increases (most noticeably at higher concentrations) while $R_{rel\%}$ decreases (most noticeably at lower concentrations.) Data for the boron method in Table 3 show this behavior. To interpret statistics that follow this model, select a procedure from Annex A2 for calculating estimates of the constants K_R and $K_{rel\%}$. Substituted in Eq 1, the constants define an equation representing the expected values of *R* for the method as a function of analyte concentration. The task group may use the relationship to estimate *R* for the method at any concentration, *C*, within the scope of the method.

$$\hat{R}_C = \sqrt{K_R^{\,2} + (C \times K_{rel\%}/100)^2} \qquad (1)$$

The boron ILS data in Table 3 yield estimates of $K_R = 0.00026$ % boron and $K_{rel\%} = 14.6$ %. The following equation predicts *R* at analyte contents from 0 % to approximately 0.012 % boron.

$$\hat{R}_C = \sqrt{0.00026^2 + (\%B \times 0.146)^2} \qquad (2)$$

7.4 *Constant Model for Error in Analytical Methods*—The ILS data follows the constant error model if, with increasing analyte concentration, *R* neither increases nor decreases but $R_{rel\%}$ continually decreases. The gold method ILS data in Table 1 show this behavior. For statistics that follow this model, use Eq 3 to calculate the root-mean-square (RMS) estimate of K_R. This value predicts *R* at all analyte contents within the scope of the method:

$$\hat{K}_{\mathrm{R}} = \sqrt{\Sigma\, R^2/n} \tag{3}$$

where:

ΣR^2 = sum of the squares of R over all test materials, and
n = number of test materials.

The six values for R, squared and added, equal 0.100901. Dividing by the number of materials, $n = 6$, and taking the square root gives an estimate for K_{R} of 0.13 %. This estimate applies from 0 to 100 % gold.

7.5 *Relative Model for Error in Analytical Methods*—The ILS data follows the relative error model, if with increasing analyte concentration, $R_{\mathrm{rel\%}}$ neither increases nor decreases but R continually increases. The manganese method ILS data in Table 2 show this behavior. For statistics that follow this model, use Eq 4 to calculate an RMS estimate of $K_{\mathrm{rel\%}}$. This value predicts $R_{\mathrm{rel\%}}$ only within the analyte content range tested during the ILS:

$$\hat{K}_{\mathrm{rel\%}} = \sqrt{\Sigma(R_{\mathrm{rel\%}})^2/n} \tag{4}$$

where:

$\Sigma(R_{\mathrm{rel\%}})^2$ = sum of the squares of $R_{\mathrm{rel\%}}$ over all test materials, and
n = number of test materials.

The six values for $R_{\mathrm{rel\%}}$, squared and added, equal 79.4161. Dividing by the number of test materials, $n = 6$, and taking the square root gives an estimate for $K_{\mathrm{rel\%}}$ of 3.7 %. This estimate applies from approximately 0.5 % to 6 % manganese.

8. Calculation of the Low Scope Limit of the Method

NOTE 2—Refer to Annex A4.

8.1 A method is always written for a nominal analyte concentration range expected to cover the anticipated applications. If the low scope limit calculated from the ILS statistics is lower than the low limit specified in the draft method, the task group need not lower the scope of the method unless it wishes to do so. However, if the calculated low limit is higher than the value specified in the draft method, the task group shall raise the low limit to the calculated value. For methods with sample portion or calibration options, calculate the low scope limit from the data for the option covering the lowest range.

8.2 The task group must establish the appropriate value for R_{L}, the lowest estimated reproducibility index from the ILS, and select e_{max}, the maximum allowable percent relative error. These constants are used to calculate L, the lowest analyte content for which the method is expected to give quantitative results.

8.2.1 R_{L}—For an ILS evaluated in accordance with the general error model (7.3) or the constant error model (7.4), set R_{L} equal to the estimate of K_{R}. Otherwise, set R_{L} equal to the estimated R for the test material with the lowest analyte content. If several have nearly equal values for R, set R_{L} equal to the square root of the sum of their squares.

8.2.2 *Maximum Allowable Error,* e_{max}—Set e_{max} equal to 50 %, a value that has been found satisfactory for most methods used to test materials at low analyte contents.

8.3 Use Eq 5 to calculate L, the lower limit as follows:

$$L = \frac{100 \times R_{\mathrm{L}}}{e_{\mathrm{max}}} \tag{5}$$

where:

R_{L} and e_{max} have the values selected in accordance with 8.2. Eq 5 becomes $L = 2\ R_{\mathrm{L}}$ for the usual case in which $e_{\mathrm{max}} = 50\ \%$.

8.4 Set the lower limit of the method to L.

9. Interpretation of ILS Statistics

9.1 A properly conducted ILS program often provides more information than is apparent from visual inspection of the statistics. Typically, when the test program is completed, the participants have recent experience with the behavior of the method as applied to different test materials. The task group may use this knowledge to clarify trends in the performance of the method at different analyte contents. If the task group agrees upon an error model that is both consistent with the observed values of R and $R_{\mathrm{rel\%}}$ and representative of their experience with the method and equipment, they may use the model to calculate the expected value for R at various analyte contents within the scope of the method as a guide to users.

9.2 *Precision Statements*—Select a statement from the following example formats:

9.2.1 *ILS in Which No Model has Been Adopted*:

[Insert the number of laboratories with data used in the ILS] laboratories participated in testing this method, providing [number of data sets actually used] sets of data. Table ____ summarizes the precision information.

9.2.2 *ILS in Which Constant Error Model has Been Adopted*:

[Insert the number of laboratories with data used in the ILS] laboratories participated in testing this method, providing [number of data sets actually used] sets of data. Table ____ summarizes the precision information. Within the scope of the method, the reproducibility index, R, is approximately [insert the estimated K_{R} from 7.4].

9.2.3 *ILS in Which Relative Error Model has Been Adopted*:

[Insert the number of laboratories with data used in the ILS] laboratories participated in testing this method, providing [number of data sets actually used] sets of data. Table ____ summarizes the precision information. Within the scope of the method, the relative reproducibility index, $R_{\mathrm{rel\%}}$ is approximately [insert the estimated $K_{\mathrm{rel\%}}$ from 7.5.]

9.2.4 *ILS in Which General Analytical Error Model has been Adopted*:

[Insert the number of laboratories with data used in the ILS] laboratories participated in testing this method, providing [number of data sets actually used] sets of data. Table ____ summarizes the precision information. The following equation predicts the approximate value of R at any concentration, C, within the scope of the method:

$$R_{\mathrm{C}} = \sqrt{K_{\mathrm{R}}^{\ 2} + \left(C \times \frac{K_{\mathrm{rel\%}}}{100}\right)^2} \tag{6}$$

Insert values for K_{R} and $K_{\mathrm{rel\%}}$ obtained in 7.3]

9.2.5 For 9.2.3 or 9.2.4, the task group may also wish to show and refer to a table of expected values for R, calculated for various analyte contents by $K_{\mathrm{rel\%}}$ or Eq 1, respectively.

9.2.6 For ILS with optional sample portions or calibration curves, the task group may need expert assistance to properly display the statistics for various concentration ranges.

9.3 *Accuracy*—The difference between the calculated mean for an analyte and its accepted value for each test material is the statistic b-value, the task group may use to judge the accuracy of a method. Include the bias information in the same table used for the precision information. Screen the b-values to eliminate any that are characteristic of the test material rather than the method. If the b-value for any material/analyte combination is much larger than for materials with similar analyte contents, look for a reason, such as poor homogeneity or a large uncertainty in the certified value. Be particularly critical of materials for which both b-value and *R* are exceptionally large. If a cause is found, remove the data for that analyte/material combination from the ILS table and recalculate the ILS statistics for that analyte. Examine each b-value separately to decide if the size of the estimated bias is so large that it creates a technical or commercial problem at the corresponding analyte content. If the task group fails to find either type of problem, the accuracy of the method shall be deemed satisfactory and no purely statistical consideration shall be advanced to question that finding.

9.4 *Bias Statements*—Select the appropriate statement from the following examples:

9.4.1 *ILS in Which No Accepted Reference Materials were Available for Testing*:

No information on the accuracy of this method is known because, at the time it was tested, no accepted reference materials were available [if other reasons apply, for example, satisfactory reference materials cannot be produced, state the applicable reason.] Users are encouraged to employ suitable reference materials, if available, to verify the accuracy of the method in their laboratories.

9.4.2 *ILS in Which One (or More) Accepted Reference Material was Tested*:

The accuracy of this method has been deemed satisfactory based upon the bias information in Table ____. Users are encouraged to use these or similar reference materials to verify that the method is performing accurately in their laboratories.

10. Keywords

10.1 error models for analytical methods; interpretation of interlaboratory studies; use of interlaboratory statistics

ANNEXES

(Mandatory Information)

A1. PRECISION MODELS FOR METHODS OF CHEMICAL ANALYSIS

A1.1 Interlaboratory studies of methods are designed to partition the total variability of the study into two or more components using analysis of variance. The total variance is broken up into the component parts revealed by the variable levels included in the statistical design. An ILS produces a separate set of statistics for each analyte in each test material (analyte/material combination.) The total variance includes all sources of variability that operate during the experiment. The statistics s_R, and *R* are the standard deviation and reproducibility index for differences between laboratories, the highest level in the ILS design. Consequently, they encompass all sources of error in the ILS.

A1.2 The total variance may also be partitioned into two mutually exclusive parts in a different manner: The first component, characteristic of the method but independent of analyte concentration, has a constant value for all test materials. The second, a function of analyte concentration, increases in materials with higher analyte contents. All constant variability sources are summed into a single variable, s_K^2, and all concentration-dependent sources into another variable, s_v^2. Because every variability source that is not constant with respect to analyte concentration must be dependent upon concentration, the total variance, s_T^2, is represented as follows:

$$s_T^2 = s_K^2 + s_V^2 \tag{A1.1}$$

Analyte concentration is implicitly included in the final term of Eq A1.1. The simplest relationship between precision of a method and analyte concentration asserts that the percent standard deviation is constant relative to concentration:

$$s_{rel\%} = 100 \times s_v/C. \tag{A1.2}$$

A1.2.1 Substitution into Eq A1.1 yields Eq A1.3, a general relationship between observed standard deviation and analyte concentration:

$$s_T^2 = s_K^2 + \left(C \times \frac{s_{rel\%}}{100}\right)^2 \tag{A1.3}$$

Note that s_K and $s_{rel\%}$ are both constants by definition. Because the reproducibility index, *R*, is a simple function of the standard deviation, *s*, Eq A1.4 follows from Eq A1.3. It represents *R* at a discrete concentration, R_C, as a function of the analyte concentration, *C*, and involves the two constants, K_R and $K_{rel\%}$, which are characteristic of the method:

$$R_C = \sqrt{K_R^2 + \left(C \times \frac{K_{rel\%}}{100}\right)^2} \tag{A1.4}$$

A1.3 *General Model for Analytical Methods*—Eq A1.4 represents a model which many ILS data sets follow. The plot of *R* for each test material in the ILS against its concentration has the following general characteristics:

A1.3.1 For analyte concentrations below a certain level, *R* becomes asymptotic to the horizontal line at $R = K_R$. The model shows K_R as the lowest possible value for *R*.

A1.3.2 For analyte concentrations above a certain level, *R* becomes asymptotic to a line through the origin with its slope equal to $K_{rel\%}/100$. The model shows $K_{rel\%}$ as the lowest possible value for $R_{rel\%}$.

A1.3.3 The unique concentration in A1.3.1 and A1.3.2 is the transition concentration, C_{trans}. At that concentration, both constant and concentration-dependent sources contribute equally to the observed variability. Eq A1.5 predicts the transition concentration:

$$C_{trans} = 100 \times \frac{K_R}{K_{rel\%}} \tag{A1.5}$$

A1.3.4 These theoretical relationships are shown in Fig. A1.1. The ILS statistics estimate the parameters of the method only at the discrete analyte concentrations of the test materials. Plots of *R* versus mean *C*, both ILS statistics, always exhibit a scatter about the fitted line. The extent of the scatter varies greatly among ILS for different methods and is affected by many factors, some of which may not be directly related to the precision of the analytical method under test. Eq A1.4 is useful because it represents the part of the data variability that can be "explained" by changes in analyte concentration between materials. Variability not related to analyte concentration is caused either by random variability in measurements or by inhomogeneity in test materials. The constants in Eq A1.4 are estimated by commonly available curve-fitting techniques, several of which are explained in Annex A2.

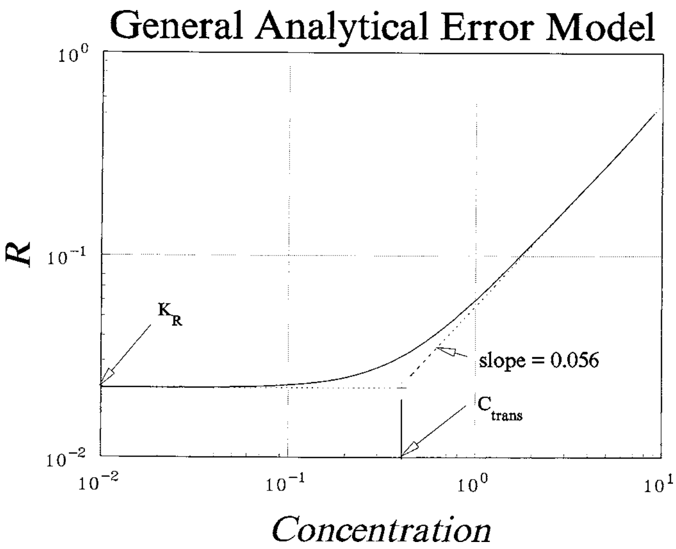

FIG. A1.1 Plot of *R* for a Hypothetical Method with $K_R = 0.022$ and $K_{rel\%} = 5.6$ %

A1.3.5 The constants of the general model have the following physical significance: K_R defines the minimum value of *R* a user may expect at low analyte contents: $K_{rel\%}$ defines the lowest $R_{rel\%}$ a user may expect for materials with higher analyte contents.

A1.4 *Constant Precision Model*—If a method is tested with materials containing less than about ½ C_{trans}, *R* will appear to be independent of analyte concentration. Some methods follow this model at all attainable analyte contents. Fire-assay methods seem to exhibit this behavior. The chemical reactions are carried out at elevated temperatures, so that the kinetics and equilibria are not significant sources of variability in comparison with weighing errors. Weighing errors are constant at all analyte contents. This explains why the ILS for a fire assay method (see Table 1) shows a constant value for *R* over the tested content range of 25 to 100 % gold. When a task group applies this model to its ILS results, it should state in the research report the chemical and physical aspects of the method causing the constant variability. This model is also implied if the estimate of $K_{rel\%}$ from the general analytical model is zero. Because *R* is independent of analyte content, the *R* value for every material is an equally probable estimate of *R*. Report the root-mean-square (RMS) average *R* of all test materials as the estimate of *R* for the method, from zero to the highest level tested in the ILS.

A1.5 *Relative Precision Model*—Many methods of chemical analysis exhibit constant relative precision over a restricted range of analyte concentrations. If the intended applications of a method lie within this range, the task group need not test the method at lower analyte levels. This model is also implied if the estimate of K_R from the general analytical model is zero. Because the relative reproducibility index is proportional to analyte concentration, the $R_{rel\%}$ of each test material is an equally probable estimate of the relative reproducibility of the method. Report the RMS average $R_{rel\%}$ of all test materials as the estimate of $R_{rel\%}$ for the method over the concentration range tested in the ILS. The lower limit of the scope is set by the requirements of the applications.

A2. PROCEDURES FOR ESTIMATING THE CONSTANTS OF THE GENERAL ANALYTICAL MODEL

A2.1 *Linear Regression Procedures*—Do not use ordinary linear regression programs, such as those provided with electronic calculators. They are not suitable for ILS data because they assume that *R* is a linear function of concentration and has constant variability over the entire range of the independent variable. In general, these assumptions are not valid for ILS data.

A2.2 *Nonlinear Curve Fitting Procedures*—Many statistical and data-plotting software packages contain procedures for fitting the nonlinear Eq A1.4 to data. The results from one program may differ from others, but the differences ordinarily are small and may usually be ignored. Always verify that the curve is a reasonable fit to the data by plotting *R* versus *C*. Use log-log coordinates to enhance visibility of the low-concentration data.

A2.3 *Least-Squares Procedures with Variable Precision*—Two calculation programs have been specifically developed to estimate the constants in Eq A1.4 for ILS data sets. The first program minimizes the squares of the differences relative to the observed values of *R*. The second minimizes the differences relative to the mean analyte concentration. The first procedure should theoretically provide the best fit, but some data sets may give more satisfactory results with the second. In the following equations, x_i represents the mean concentration, C_i, of a test material, and y_i represents R_i (observed R) for the same material. The number of test materials is m. The summations are over all m data pairs.

A2.3.1 *Precision Proportional to R*—The following equations estimate the model constants by minimizing the sum of $(R_i^2 - R_{calc}^2)/R_i^2$ for all test materials:

$$A^2 = \frac{m\Sigma\frac{x_i^4}{y_i^2} - \Sigma x_i^2\,\Sigma\frac{x_i^2}{y_i^2}}{D_1} \qquad (A2.1)$$

If A^2 is negative, set $\hat{K}_R = -\sqrt{|A^2|}$, otherwise $\hat{K}_R = \sqrt{A^2}$

$$B^2 = \frac{\Sigma x_i^2\frac{1}{y_i^2} - m\,\Sigma\frac{x_i^2}{y_i^2}}{D_1} \qquad (A2.2)$$

If B^2 is negative, set $\hat{K}_{rel} = -\sqrt{|B^2|}$, otherwise $\hat{K}_{rel} = \sqrt{B^2}$;

where:

$$D_1 = \Sigma\frac{1}{y_i^2}\Sigma\frac{x_i^4}{y_i^2} - \left(\Sigma\frac{x_i^2}{y_i^2}\right)^2 \qquad (A2.3)$$

A2.3.2 *Precision Proportional to Concentration*—The following equations estimate the model constants by minimizing the sum of $(R_i^2 - R_{calc}^2)/C_i^2$ for all test materials:

$$A^2 = \frac{\Sigma x_i^2\,\Sigma\frac{y_i^2}{x_i^2} - m\,\Sigma y_i^2}{D_2} \qquad (A2.4)$$

If A^2 is negative, set $\hat{K}_R = -\sqrt{|A^2|}$, otherwise $\hat{K}_R = \sqrt{A^2}$;

$$B^2 = \frac{\Sigma y_i^2\,\Sigma\frac{1}{x_i^2} - m\,\Sigma\frac{y_i^2}{x_i^2}}{D_2} \qquad (A2.5)$$

If B^2 is negative, set $\hat{K}_{rel} = -\sqrt{|B^2|}$, otherwise $\hat{K}_{rel} = \sqrt{B^2}$;

where:

$$D_2 = \Sigma x_i^2\,\Sigma\frac{1}{x_i^2} - m^2 \qquad (A2.6)$$

A2.3.3 *Example 1*—The boron in steel ILS statistics from Table 3 are used to illustrate the three procedures for estimating the constants of Eq A1.4. These ILS data are unusually extensive, the combined data from two studies. One involved 8 materials and 14 laboratories, and the second 8 materials and 21 laboratories. The data are combined because a previous study demonstrated that each ILS independently provided estimates of the same method statistics. The uppermost curve was obtained by means of a commonly used nonlinear fitting program that uses the Marquardt-Levenberg least-squares method. Eq A2.4 and Eq A2.5 produced the next lower curve (barely distinguishable from the first.) Eq A2.1 and Eq A2.2 produced the bottom curve. Although there is little difference among the three predictive equations, the lowest curve seems the most satisfactory. The equation predicting R is $R_C = \sqrt{(0.000216)^2 + (14.51 \times C/100)^2}$. Predicted reproducibilities at various analyte contents are shown in Table A2.1. The curve shows that R probably exceeds 50 % of the mean boron result, for materials containing less than approximately 0.0005 % boron. The lower scope limit, L, is $(2 \times 0.000216) = 0.00043$, which the task group should round to the next higher significant digit, 0.0005 %.

TABLE A2.1 Predicted Reproducibility Index for Boron

Boron, %	Reproducibility Index, *R*
0.0001	0.00022
0.0005	0.00023
0.001	0.00026
0.003	0.00049
0.006	0.00090
0.009	0.00132
0.012	0.00175

TABLE A2.2 Reproducibility Index of Iron in Refined Gold

Material	Found, ppm	Reproducibility Index, *R*	$R_{rel\%}$
3A	2.4	1.22	51
1A	4.1	1.53	37
2A	4.3	1.31	30
4A	141.3	6.81	4.82

A2.3.4 *Example 2*—Statistical information for an ILS of a method for iron in refined gold is shown in Table A2.2. Seven laboratories analyzed four samples. This is a very sparse study, but with the aid of the analytical precision model, the task group provides the user with estimates of R for the method from 0 to about 150 ppm. Three of the materials clustered at the low end of the expected range for the method, which gives a good estimate of the minimum value for R, 1.34 ppm. Fig. A2.2 shows how well the model fits the data. The prediction equation is $r = \sqrt{1.34^2 + (\text{ppm Fe} \times 0.0473)^2}$. Table A2.3 is an example of the reproducibility information the task group provides using this equation.

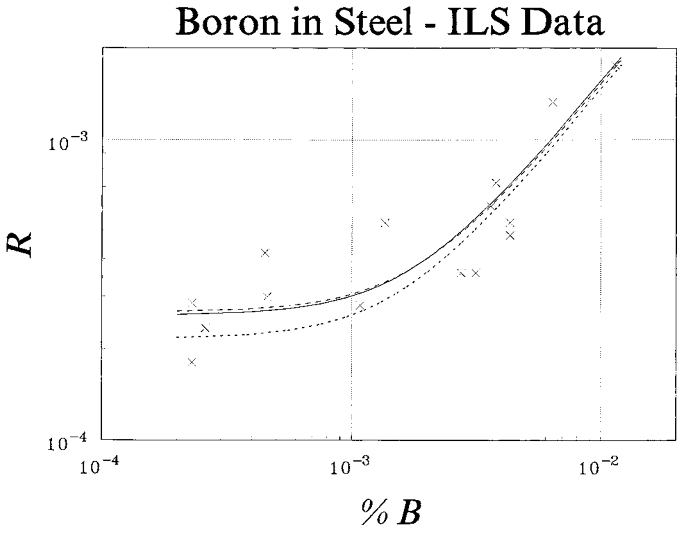

FIG. A2.1 Variation of Reproducibility Index with Analyte Concentration—General Model

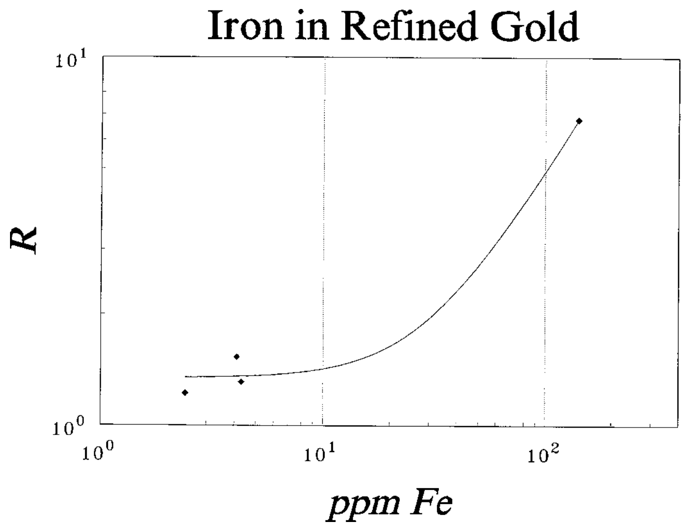

FIG. A2.2 *R* for Iron in Refined Gold

TABLE A2.3 Predicted Reproducibility Index of Iron in Refined Gold

Iron, ppm	Reproducibility Index, *R*	Iron, ppm	Reproducibility Index, *R*
5	1.4	90	4.5
20	1.6	125	6.1
50	2.7	150	7.2

A3. FACTORS AFFECTING THE USEFULNESS OF ILS DATA FOR EVALUATING METHODS

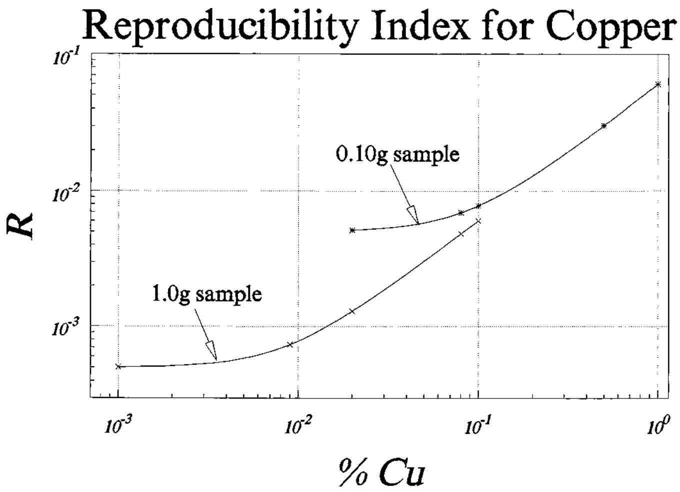

FIG. A3.1 Hypothetical Copper Method

A3.1 *Special Requirements for Methods with Extended Ranges*—Some methods use measurement processes that apply to a limited range of analyte concentrations. For applications that require coverage of a large range, these methods use larger or smaller sample portions to decrease or increase the normal concentration range of the method. The sensitivity of certain instruments may be altered to provide different calibrated ranges, as required by the analyte content of samples. A method is likely to exhibit different precision characteristics when either sample size or calibration range is varied. If a method allows either type of modification, it must be tested as if each specified sample size or each specified calibrated range were a separate method.

A3.1.1 *Example 1*—A hypothetical copper method has a calibrated range of 0 to 0.0100 µg/mL. It covers from 0 to 0.1 % Cu for 1 g samples and 0 to 1 % Cu for 0.1 g samples. The sample solution is diluted to 100 mL. Test materials having copper contents of 0.001, 0.009, 0.02, 0.08, 0.10, 0.50, and 1.0 % Cu are used in the ILS. The five lowest content materials are analyzed with 1 g samples, and the highest five with 0.1 g samples. This is accomplished by using three materials at both levels. Assume that $K_R = 0.0005$ and $K_{rel\%} = 6.0\ \%$. The factors for converting solution concentration to percent copper in the materials are 10 for 1 g samples and 100 for 0.1 g samples. Also, *R* is multiplied by the same factor, giving the trends for *R* shown in Fig. A3.1. The equations are as follows:

$$R_{1.0} = 10 \times \sqrt{(0.0005)^2 + \left(\frac{C_s}{10} \times 0.06\right)^2} \text{ and} \tag{A3.1}$$

$$R_{0.1} = 100 \times \sqrt{(0.0005)^2 + \left(\frac{C_s}{100} \times 0.06\right)^2} \tag{A3.2}$$

where C_s = analyte content of the solid test sample.

In this example, the chemistry of the color formation is unaffected by sample dilution. One calibration curve is valid for both sample weights. This example illustrates why higher ranges of multiple-range methods must be tested with at least one material from the next lower range to properly characterize the precision performance of the method in each range.

A3.1.2 *Example 2*—Table A3.1 shows data for an ILS in which sample dilution affects the precision. Separate calibration curves are required for each sample dilution. The constants of the analytical model are different for each sample size. The method options using smaller test portions were not tested with materials at lower contents, so observed values for *R* at the low end of the higher ranges were not obtained in the ILS, as Fig. A3.2 reveals. Table A3.2 shows one way the results of the ILS may be summarized for inclusion in a test method.

A3.1.3 *Example 3*—Table A3.3 and Fig. A3.3 illustrate an ILS in which an instrument is calibrated over different analyte content ranges. The gaps in the *R* information are apparent. As with methods having test-portion options, this ILS could have provided complete coverage by including a lower material in the test of each of the higher ranges. The equations for the analytical model were estimated, although the constants, especially K_R, are not as well characterized as they could have been with the suggested additional data. By range they are as follows:

$$0.01\ \%, R = \sqrt{0.000922^2 + (\%C \times 0.2556)^2}$$

$$0.1\ \%,\ \ R = \sqrt{0.004113^2 + (\%C \times 0.05418)^2}$$

$$1.0\ \%, R = \sqrt{0.009456^2 + (\%C \times 0.04015)^2}$$

$$4.5\ \%,\ \ R = \sqrt{0.01911^2 + (\%C \times 0.02457)^2}$$

These equations produced the values of *R* in Table A3.4. Note that the last column shows a lower predicted value for *R* for the 4.5 % range between 0.5 and 1.0 % *C* than for the 1.0 % range. Perhaps the original choice of ranges was not optimum to get the best precision from the method at all analyte contents.

A3.2 *Test Materials, Distribution of Analyte Contents*—The task group usually chooses materials from a restricted list of those appropriate for the test method. It is not always possible, but the task group should attempt to include at least two materials with analyte content high in each optional range and another at a low level. Fig. A3.1 shows that seven test materials are sufficient to test two ranges at five levels in each range. The ILS shown in Fig. A3.2 and Fig. A3.3 could have met this criterion by using low content materials already included in the study for testing the higher ranges. For the lowest range in each method, include several test materials (which do not have to be accepted reference materials) at or below the lowest analyte level of interest. Ideally, at least two more test materials are required, one within about 40 to 75 %, and the other within 90 to 100 % of highest level in the calibrated range.

TABLE A3.1 Copper in Iron by the Neocuproine Photometric Method

	Sample Weight, g	Copper found, %	Reproducibility Index, *R*	$R_{rel\%}$
1	0.2	0.006	0.004	66.7
2	0.2	0.014	0.006	42.9
3	0.2	0.033	0.004	12.1
4	0.2	0.078	0.010	12.8
5	0.2	0.118	0.016	13.6
6	0.2	0.176	0.021	11.9
7	0.06	0.200	0.018	9.0
8	0.06	0.221	0.022	10.0
9	0.06	0.361	0.036	10.0
10	0.01	1.51	0.05	3.31
11	0.01	5.53	0.18	3.25

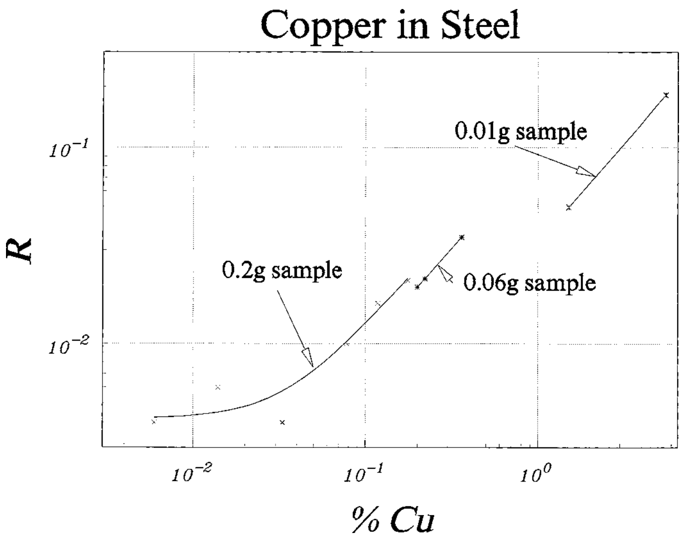

FIG. A3.2 Neocuproine Copper Method

TABLE A3.2 Predicted Reproducibility Index for Copper in Iron

Test Portion, g	Copper Range, %	Reproducibility Index, *R*	$R_{rel\%}$
0.2	0–0.03	0.0047	...
0.2	0.03–0.20	...	12.6
0.06	0.20–0.60	...	9.7
0.01 or less	0.6–7.5	...	3.3

TABLE A3.3 Statistical Information for Carbon, Combustion/IR

	Calibration Range, % C	Certified, % C	Found, %C	Reproducibility Index, *R*	$R_{rel\%}$
1	0 to 0.01	<0.002	0.0013	0.00105	80.8
2		0.0047	0.0044	0.0013	29.5
3		0.0068	0.0075	0.0022	29.3
4	0 to 0.10	0.026	0.0257	0.0044	17.1
5		0.042	0.0409	0.0046	11.2
6		0.094	0.0926	0.0065	7.02
7	0 to 1.0	0.18	0.182	0.012	6.59
8		0.452	0.456	0.0205	4.50
9		0.753	0.760	0.032	4.21
10	0 to 4.5	1.26	1.266	0.039	3.08
11		2.54	2.559	0.059	2.30
12		4.12	4.095	0.105	2.59

A3.3 *Quality of Test Materials*—Of the three factors involved in an ILS, the test materials are considered the most reliable. Methods are expected to exhibit variability. Individual laboratory failures are obvious when results are compared among laboratories. Few have thought to challenge the integrity of the test materials. Most reference materials deserve our confidence, but the task group must be aware that some types

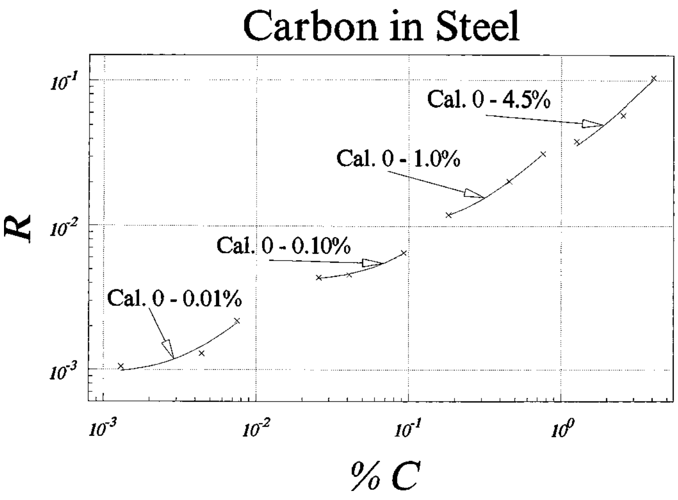

FIG. A3.3 Carbon, Combustion/IR Method

TABLE A3.4 Predicted Reproducibility Index for Carbon, Combustion/IR

Carbon Content, % C	Reproducibility Index, 0.01 % Range	Reproducibility Index, 0.10 % Range	Reproducibility Index, 1.0 % Range	Reproducibility Index, 4.5 % Range
0.001	0.00096			
0.003	0.00120			
0.005	0.00158	0.00412		
0.010	0.00272	0.00415		
0.02		0.00425		
0.05		0.00493	0.00968	
0.10		0.00680	0.01028	
0.2			0.0124	
0.5			0.0222	0.0227
1.0			0.0413	0.0311
2.0				0.0527
4.5				0.1122

TABLE A3.5 Reproducibility Index Statistics for Sulfur in Iron and Steel

Material	Sulfur Found, %	Reproducibility Index, *R*	$R_{rel\%}$
1	0.1213	0.01713	14.1
2	0.1475	0.02282	15.5
3	0.1480	0.01945	13.2
4	0.1547	0.02304	14.9
5	0.2004	0.01856	9.76
6	0.2398	0.02883	12.0
7	0.2913	0.02567	8.81
8	0.3383	0.02251	6.65
D1	0.2913	0.02567	8.81
D2	0.2330	0.05443	10.5
D3	0.1756	0.01880	10.7
D4	0.1182	0.01510	12.8
D5	0.05958	0.01136	19.1

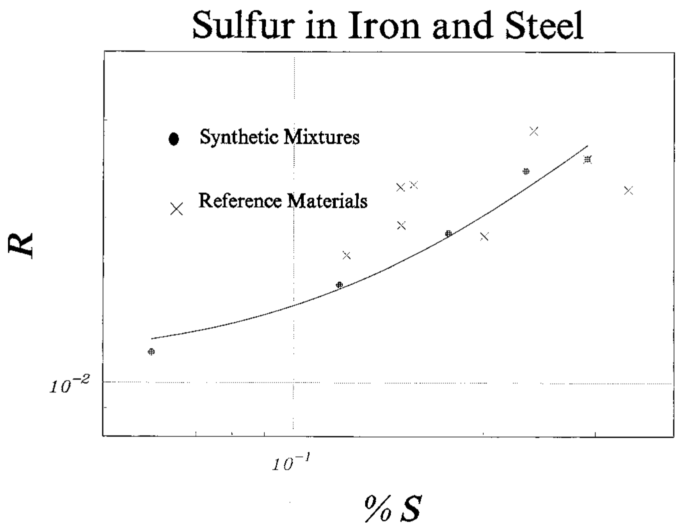

FIG. A3.4 Sulfur in Iron and Steel

of materials are inherently difficult to prepare in a homogeneous condition. Occasionally, the number on a certificate is not correct.

A3.3.1 *Homogeneity of ILS Test Materials*—Most solid-form test materials have some degree of heterogeneity even when prepared as chips, granules, or powders. Only when a material is completely dissolved and mixed in a true solution, is it unquestionably homogeneous. Reliable suppliers of reference materials usually test their products to ensure that their heterogeneity is small. The task group cannot ignore the possibility that a specific material or type of material may be variable enough in composition to affect the ILS statistics.

A3.3.1.1 *Example 4*—Table A3.5 shows the ILS results from a method for high sulfur in iron and steel. Two types of test materials were included in the study. Materials 1 through 8 were produced by casting solids of the desired approximate sulfur content and reducing them to fine particles. The second set, marked D1 through D5, were produced by using a large portion of Sample 7 as the base material for dilution with increasing amounts of low-sulfur iron powder. The task group expected this series of mixtures to have known sulfur contents derived from the base material. Table A3.5 and Fig. A3.4 shows this to be the case. However, the observed *R* values for the other reference materials are probably characteristic of the individual materials, rather than the test method. The model equation for the mixtures data is as follows:

$$R = \sqrt{0.01087^2 + (\%S \times 0.08546)^2} \tag{A3.3}$$

It is not surprising that $K_{rel\%}$ for the model, 8.5 %, is about the same as for Material 7 (8.8 %), the base material from which the mixtures were prepared. It would be a mistake to claim that this prediction equation is characteristic of the method. Instead, the equation represents some not-too-meaningful combination of the variability of the method with the compositional variation within the specific synthetic mixtures. For the method applied to homogeneous materials *R* would probably be smaller than the curve suggests. The task group may accept the method, even though an unequivocal precision statement is not appropriate. A statement concerning the difficulty in obtaining representative homogeneous samples would be helpful to users.

A3.3.1.2 *Example 5*—Table A3.6 shows ILS results for a method of determining tantalum in various high-temperature alloys. Four reference materials with certified values and one low material with a non-certified value were included. A question raised about the relatively large bias for Material 3 implied that the method was inaccurate. The task group believed the method was sufficiently accurate for its proposed

TABLE A3.6 Statistical Information for Tantalum in High Temperature Alloys

Material	Certified Ta, %	Found Ta, %	Bias, %	Reproducibility Index, R	$R_{rel\%}$
1	(0.012)	0.0099	...	0.0077	78
2	0.027	0.0275	0.005	0.0092	33
3	0.18	0.168	−0.012	0.083	49
4	0.34	0.348	0.008	0.099	28
5	0.95	0.944	−0.006	0.206	22

uses. The task group noticed that $K_{rel\%}$ for Material 3 was also larger than expected in comparison with other test materials. The certification data for the material was investigated and cause was found to eliminate Material 3 from the ILS. The dotted line in Fig. A3.5 shows the great effect of one poor reference material on an ILS carried out with few test materials. When Material 3 was eliminated, the model provided a good fit to the remaining materials (solid line.)

A3.4 *Performance of Individual Laboratories*—The ILS is a test of a method, not of laboratories. Just as test materials may occasionally not perform as they should, individual laboratories may also add variability that should not be attributed to the method.

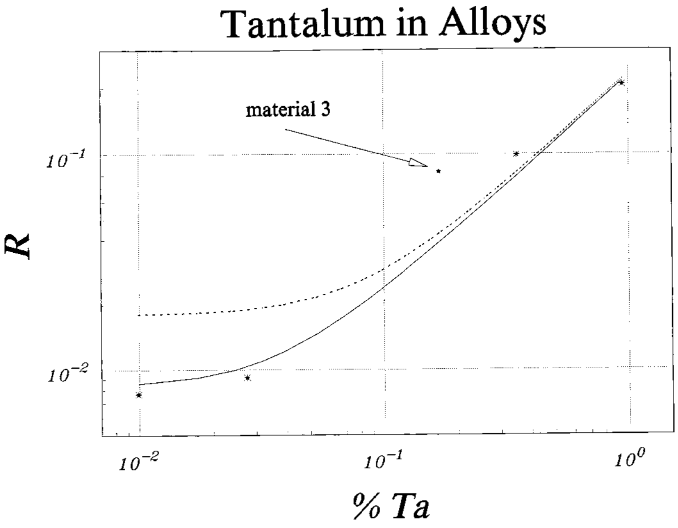

FIG. A3.5 Tantalum in Alloys

A3.4.1 Inexperience or ineptitude may cause an analyst to improperly perform an operation that is clearly explained in the text. If the laboratory cannot provide qualified personnel or does not have the necessary equipment, it must be dropped from the study. Even highly experienced analysts have an occasional lapse, and if a specific mistake can be identified and eliminated, the laboratory should be encouraged to resubmit its contribution to the study. When wording in the method contributes to misinterpretation of the way to perform an operation, the text must be edited carefully until instructions are clear.

A3.4.2 A problem is created during the laboratory testing phase of an ILS by contributors who "know a better way to perform the method," or who insist on doing it as it is usually done in that laboratory. The task group must not tolerate such irresponsible behavior among contributing laboratories. It is unfair to the method and to the other laboratories taking part in the study. If the innovation is a significant one, it can only increase the variability exhibited by the ILS. If it is not a significant change, then it may become an unnecessary complicating factor in an already complex task.

A3.4.3 Task groups should be aware that the method text, in a form ready for testing, shall be made available to all prospective contributing laboratories so that they may have time to examine it before the test begins. Giving the task group members an opportunity to comment, make suggestions, and question what is meant by wording that may be unclear to them may produce a better method and always avoids problems during testing. Valid suggestions at this stage may result in changes that are easily made in the method before the laboratory test phase of the ILS begins. However, after all discussion is over, the task group must adopt an official test version of the method and all participants agree to perform the method exactly as it is written.

A4. APPLICATIONS OF QUANTITATION AND DETECTION PROCESSES

A4.1 *Statistical Basis for Quantitation*—In methods for chemical analysis, the term "quantitation" refers to the ability of the method to produce results having a defined risk of exceeding a specified maximum relative error in interlaboratory comparisons. The purpose is to ensure that the method yields meaningful numerical values. To be designated "quantitative," the scope of the method shall specify the range of analyte contents within which the method meets selected "quantitation criteria." The most critical factor is the low limit of that range. A user attempting to analyze materials with lower analyte content risks reporting results that may not be verifiable in other laboratories. The usual quantitation criteria are as follows:

A4.1.1 In conformance with Practice E1601, a task group shall use statistics having a 95 % confidence level. The appropriate statistic is R, obtained at approximately the required 95 % confidence level from the ILS.

A4.1.2 The task group shall also establish the maximum acceptable relative difference between results obtained in different laboratories. An allowable maximum relative error, e_{max}, of 50 % for interlaboratory differences has been found generally satisfactory for methods used to determine trace contaminants in metals, alloys, ores, and related materials. This value shall be used unless the task group specifically justifies a lower value based upon the applications of the method.

A4.1.3 *Defining Equation for Low Limit*—At any concentration, R gives a measure of the difference between results obtained on the same material in two laboratories with a 95 % probability. In general, the observed difference will be smaller than R for 19 of 20 comparisons. This means that the relative percentage error will be less than (100 × *R/C*) at concentration *C* with the same probability. Rearranged and with *R* set to the value of K_R, this equation is useful for calculating the lowest analyte content at which the quantitation criteria will be met:

$$C_{QL} = \frac{100 \times K_R}{e_{max}} \qquad (A4.1)$$

This concentration, C_{QL}, is the desired value for the low scope limit. An attempt to analyze materials having a lower analyte content will produce results with greater error than the allowable limit at the 95 % confidence level.

A4.2 *Statistical Basis for Detection*—"Detection" is an event in which the mean response for an analyte on a specific instrument exceeds the statistically defined "detection limit" for that analyte and instrument. Various definitions have been proposed for detection limit, but they differ only in the confidence level specified. Invariably, the detection limit is calculated from standard deviations of repeated measurements of the response of an instrument in the absence of, or at low concentrations of, an analyte. Standard deviations from several instruments may be combined to give a typical detection limit for that analyte and that type of instrument. In any case, a detection limit includes only variability from sources that operate during the interval in which the repeat readings for the standard deviation are collected. In statistical terms, it would be described as a single-instrument, single-operator standard deviation obtained under repeatability conditions. If many such standard deviations are combined (pooled) into a single value, the result is still nothing more than a pooled single-instrument, single-operator standard deviation obtained under repeatability conditions. It does not include longer-term sources of variability within or between laboratories such as differences in calibration standards, recalibration, or changes in the analyst or the condition of instruments or equipment. Because the latter are significant components of the total error observed in the ILS, the detection limit is not a suitable statistic for judging method performance. To be useful in judging the performance of a method, standard deviations shall be obtained under reproducibility conditions to include all sources of variability that act between laboratories.

A4.3 A task group shall not use detection limits as the basis for setting the lower scope limit of a method. It shall not in any way state or imply that the detection limit is a measure of the precision of a method, unless it has convincing evidence that between-laboratory and long-term within-laboratory sources of variability are small enough to be ignored.

APPENDIXES

(Nonmandatory Information)

X1. RATIONALE

X1.1 *Approach*—This guide is based upon analytical chemistry and draws upon the basic disciplines of chemistry, physics, mathematics, and statistics. It is based upon the knowledge and experience accumulated by analytical chemists in the practical study of techniques and procedures for determining the composition of materials. The focus is on the characteristics and behavior of analytical methods and ways to provide users with the information necessary to understand them.

X1.1.1 Methods have a scope to describe applications the method serves and a section titled "Precision and Bias" to describe the attributes of precision and accuracy. Precision is the degree to which the conditions of the method may be recreated from time to time and from one laboratory to another. Accuracy is the fidelity with which the method responds to different analyte levels and material types. Precision and accuracy are attributes because they do not have explicit numerical values, yet they are useful for comparative purposes under circumstances when quantification of the associated statistics is difficult or impossible. For example, we might expect an untried new method to have precision and accuracy similar to those of a well-characterized method that uses the same principles and techniques. The terms are also to be preferred when discussing the error behavior of a method in generalized expressions rather than as discrete statistical values.

X1.1.2 Theoretical expectations are not always fulfilled in practice. For this reason, methods are tested by analyzing either synthetic or real samples of known composition. Chemists believe that some methods exhibit variability proportional to the analyte concentration, while others have an error that is constant. Still others exhibit both modes of behavior over their range of application. Until recently, the kind of data necessary to study the behavior of a wide sampling of analytical methods and techniques was not available. The development of the interlaboratory study (ILS) by ASTM and other standards organizations has produced a quantity of variability data for many different kinds of methods. Study of the patterns of precision exhibited by these methods over wide ranges of analyte concentration has revealed a relationship that is remarkably consistent. This study developed a mathematical model that will help task groups to understand more clearly the meaning of the limited data they obtain from an ILS.

X1.2 *Basis for General Model for Analytical Precision*—Inspection of ILS data sets revealed that the two basic error trends noted by analytical chemists usually were observed in the data of nearly every method tested over a sufficiently wide concentration range. The variability of results on a material containing an insignificant quantity of the analyte does not equal zero. If readings are taken with sufficient care, a finite standard deviation is always observed for a zero result. For a given method, the standard deviation remains approximately the same, even for analytes at low but measurable levels. As the analyte content is further increased, the variability begins to rise, until at elevated concentrations, the rate of change eventually becomes and remains proportional to the increase in analyte content. This behavior suggests that variability of results for each method may be described by two components. One is constant for a method (independent of analyte concentration.) The other is zero when no analyte is present but increases with increasing analyte content. The model relates the variability of results at any given concentration to the sum of the variabilities of the two components. If variability is measured as a standard deviation, the sum of the squares of the components equals the square of the observed standard deviation. The reproducibility index, *R*, follows the same model.

X1.3 *Role of ILS Statistics in Method Evaluation*—An ILS provides estimates of a method's variability only at the analyte contents and in the matrix types of the test materials used in the test program. Even the most extensive ILS provides meager coverage, a few materials representing analyte levels from zero to the highest permitted in the scope of the method. If the method applies to more than one type of material, the problem is compounded because a practical test may include only a few of the infinite number of compositions possible within each matrix and not even include each distinct matrix type.

X1.3.1 The ILS is obviously a very sparse sampling of the materials a method may cover. Each ILS is applied to a single test material. A statistical test helps task groups identify mistakes made by a participating laboratory, but most ILS programs are unable to discriminate between variability caused by material inhomogeneity and variability associated with the method. Lacking a direct statistical test, a task group must choose its test materials with care. Even then, an occasional not-quite-homogeneous material may unknowingly be included which distorts the observed precision at its concentration.

X1.3.2 If precision of methods followed no laws or if each method followed unpredictable laws, ILS statistics would be useless. The standard deviation for results would be predictable only at the discrete concentrations and in the specific matrices represented by the test materials. If the variability in results changed with concentration in an arbitrary way, it would be impossible to evaluate the precision of any method. The quantity of data provided by an ILS is insufficient to distinguish among the many possible mathematical models unless prior knowledge permits us to predict the error behavior in the gaps between test materials. ILS statistics are not useless because we observe them following simple trends. If *R* is plotted against the concentration, the curve exhibits a generally smooth upward trend. This does not mean that the proper model for analytical precision is strictly linear. The hyperbolic model regularly provides a better fit to ILS statistics than any linear model. The inspiration for the hyperbolic precision model came from many sets of statistics from the ILS of many methods. Though imperfect in detail and completeness as separate sets of statistics, taken together, they confirm the validity of the model.

X1.3.3 The ILS statistics for a test material (*R* and mean) are not parameters of a method. They are parameters of a data set produced by the operation of that method on that test material. The task group decides whether each such set of statistics is more associated with the performance of the method or with variability in the material. Once accepted, the model becomes a powerful tool to help the task group extract meaning from its ILS statistics. Because the model represents the way the precision attribute operates in the method, each "normal" ILS statistic makes its proper contribution to the estimates of the constants that define the method's precision curve. If *R* for a specific test material includes extraneous variability because of compositional inhomogeneity, that point falls above the curve defined by the more homogeneous materials.

X2. EQUATIONS FOR LEAST-SQUARE ESTIMATION OF MODEL CONSTANTS

X2.1 *Procedure*—The constants in the ordinary linear least-square program for fitting a straight line to data pairs minimize the sum of the squares of the differences between observed and calculated *y* at each given value of *x*. This familiar model assumes that the variance is constant across all values of the independent variable in the model. The general model for analytical precision is nonlinear; it is a second-order equation that produces a graph resembling half a hyperbola. The same mathematical procedures used to derive equations for the slope and intercept of linear equations can be used to derive equations for the constants in the hyperbolic equation, the constant reproducibility index, K_R, and relative reproducibility index, K_{rel}.

X2.1.1 Two-curve fitting procedures are to be derived. The least-square differences may be minimized as a function of concentration, *C*, or as a function of the observed *R*. For each of the *m* experimental points (data pairs *C*, *R*) the residual r_i represents the difference between the calculated reproducibility index, R_{Ci}, and the observed reproducibility index, R_{Oi}:

$$r_i^2 = R_{Ci}^2 - R_{Oi}^2 \qquad \text{(X2.1)}$$

The least-square curve fitting criterion requires minimizing the sum of the squares of the relative residuals, S_r. The latter is r_i divided by either R_{Oi} or by C_i:

$$S_r = \sum_{i=1}^{m} \left(\frac{r_i^2}{R_{Oi}^2} \right)^2 = \textit{minimum} \qquad \textbf{r relative to R} \qquad \text{(X2.2)}$$

$$S_r = \sum_{i=1}^{m} \left(\frac{r_i^2}{C_i^2} \right)^2 = \textit{minimum} \qquad \textbf{r relative to C} \qquad \text{(X2.3)}$$

X2.1.2 The equation used for fitting the curves to the data is the general model for analytical precision:

$$R_{Oi}^2 = K_R^2 + C_i^2 \times K_{rel}^2 \qquad \text{(X2.4)}$$

Combining Eq X2.1 with Eq X2.4 yields a general equation for the least-square process, which leads to the two desired defining equations needed to fit the curve to data sets:

$$r_i^2 = K_R^2 + C_i^2 \times K_{rel}^2 - R_{Oi}^2 \tag{X2.5}$$

$$S_r = \sum_{i=1}^{m} \left(\frac{K_R^2 + C_i^2 \times K_{rel}^2 - R_{Oi}^2}{R_{Oi}^2} \right)^2 \quad \textbf{relative to reproducibility} \tag{X2.6}$$

$$S_r = \sum_{i=1}^{m} \left(\frac{K_R^2 + C_i^2 \times K_{rel}^2 - R_{Oi}^2}{C_i^2} \right)^2 \quad \textbf{relative to concentration} \tag{X2.7}$$

Each equation is differentiated with respect to K_R and K_{rel} and the result set equal to zero. These pairs of simultaneous equations are then solved for the constants for each case.

X2.1.3 Minimum error relative to R is obtained with constants derived from Eq X2.6. To simplify the equations, let $A = K_R$ and $B = K_{rel}$:

$$\frac{dS_r}{dA} = \Sigma \frac{1}{R_{Oi}^2} A^2 + \Sigma \frac{C_i^2}{R_{Oi}^2} B^2 - m \tag{X2.8}$$

$$\frac{dS_r}{dB} = \Sigma \frac{C_i^2}{R_{Oi}^2} A^2 + \Sigma \frac{C_i^4}{R_{Oi}^2} B^2 - \Sigma C_i^2 \tag{X2.9}$$

Eq X2.8 and Eq X2.9 are simultaneous equations in A^2 and B^2 and may be solved by matrix manipulations:

$$\begin{bmatrix} \Sigma \frac{1}{R_{Oi}^2} & \Sigma \frac{C_i^2}{R_{Oi}^2} & m \\ \Sigma \frac{C_i^2}{R_{Oi}^2} & \Sigma \frac{C_i^4}{R_{Oi}^2} & C_i^2 \end{bmatrix} \quad \textbf{Solution Matrix} \tag{X2.10}$$

$$D_1 = \Sigma \frac{1}{R_{Oi}^2} \Sigma \frac{C_i^4}{R_{Oi}^2} - \left(\Sigma \frac{C_i^2}{R_{Oi}^2} \right)^2 \quad \textbf{Denominator} \tag{X2.11}$$

$$K_R^2 = A^2 = \frac{m \Sigma \frac{C_i^4}{R_{Oi}^2} - \Sigma C_i^2 \Sigma \frac{C_i^2}{R_{Oi}^2}}{D_1} \tag{X2.12}$$

$$K_{rel}^2 = B^2 = \frac{\Sigma C_i^2 \Sigma \frac{1}{R_{Oi}^2} - m \Sigma \frac{C_i^2}{R_{Oi}^2}}{D_1} \tag{X2.13}$$

X2.1.4 Minimum error relative to C is obtained with constants derived from Eq X2.7. To simplify the equations, Let $A = K_R$ and $B = K_{rel}$:

$$\frac{dS_r}{dA} = \Sigma \frac{1}{C_i^2} A^2 + mB^2 - \Sigma \frac{R_{Oi}^2}{C_i^2} \tag{X2.14}$$

$$\frac{dS_r}{dB} = mA^2 + \Sigma C_i^2 B^2 - \Sigma R_{Oi}^2 \tag{X2.15}$$

Eq X2.14 and Eq X2.15 are simultaneous equations in A^2 and B^2 and may be solved by matrix manipulations:

$$\begin{bmatrix} \Sigma \frac{1}{C_i^2} & m & \Sigma \frac{R_{Oi}^2}{C_i^2} \\ m & \Sigma C_i^2 & \Sigma R_{Oi}^2 \end{bmatrix} \quad \textbf{Solution Matrix} \tag{X2.16}$$

$$D_2 = \Sigma C_i^2 \Sigma \frac{1}{C_i^2} - m^2 \quad \textbf{Denominator} \tag{X2.17}$$

$$K_R^2 = A^2 = \frac{\Sigma C_i^2 \Sigma \frac{R_{Oi}^2}{C_i^2} - m \Sigma R_{Oi}^2}{D_2} \tag{X2.18}$$

$$K_{rel}^2 = B^2 = \frac{\Sigma R_{Oi}^2 \Sigma \frac{1}{C_i^2} - m \Sigma \frac{R_{Oi}^2}{C_i^2}}{D_2} \tag{X2.19}$$

X2.1.5 For certain sets of ILS data, A^2 and B^2 may evaluate to negative quantities. For the analytical error model, a negative value for K_R or K_{rel} has no physical significance. It indicates that some aspect of the ILS was flawed. The task group should attempt to identify and correct the causes of this behavior and reevaluate the statistics for the method.