Designation: E 1914 – 98 (Reapproved 2003)

Standard Practice for Use of Terms Relating to the Development and Evaluation of Methods for Chemical Analysis[1]

This standard is issued under the fixed designation E 1914; 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.

INTRODUCTION

ASTM methods for determining the chemical composition of materials usually are developed in four stages: (*1*) experimental development of procedures and techniques, (*2*) translation of research into text suitable for analysts (in ASTM format), (*3*) demonstration of performance in an interlaboratory study (ILS), and (*4*) acceptance as a method published for use in laboratories. Details of the development processes may be complex, but the common concepts and terms needed to discuss them are relatively simple. The concepts must be carefully defined and terms selected to represent them unambiguously in the intended contexts.

A list of terms and definitions does not guarantee clear communication. Many terms have different common and technical meanings while representing different concepts when used in various contexts. The use of important terms and concepts in the context of methods of chemical analysis is illustrated by descriptions and by examples to help task group and subcommittee members communicate clearly.

1. Scope

1.1 This document covers terms and concepts used in developing and evaluating the performance of methods for determining chemical composition. Although useful with many types of methods, they are dealt with in this document in the context of chemical analysis of metals and related materials.

[1] This practice is under the jurisdiction of ASTM Committee ~~E-1~~ E01 on Analytical Chemistry for Metals, Ores and Related Materials and is the direct responsibility of Subcommittee E01.22 on Statistics and Quality Control.

Current edition approved ~~May 10, 1998.~~ Oct. 1, 2003. Published ~~July 1998.~~ November 2003. Originally ~~published as E 1914–97.~~ approved in 1997. Last previous edition approved in 1998 as E 1914 – 97.

2. Referenced Documents

2.1 *ASTM Standards:* [2]

E 135 Terminology Relating to Analytical Chemistry for Metals, Ores, and Related Materials

E 1601 Practice for Conducting an Interlaboratory Study to Evaluate the Performance of an Analytical Method

E 1763 Guide for Interpretation and Use of Results from Interlaboratory Testing of Chemical Analysis Methods

3. Terminology

3.1 *Definitions Relating to Analytical Methods:*

3.1.1 *accuracy, n—of methods of chemical analysis,* a characteristic manifested by agreement between average results and true analyte contents.

3.1.2 *analyte, n—in methods of chemical analysis,* the component determined by a method.

3.1.3 *matrix, n—in methods of chemical analysis,* all components in a material except the analyte.

3.1.4 *method, n*—instructions used to produce a numerical result which are detailed in a document also referred to as "the method."

3.1.5 *precision, n—of methods of chemical analysis,* a characteristic manifested by agreement among individual results at a given analyte content.

3.1.6 *result, n*—value representing the quantity of analyte that is obtained by applying a method one time to a test material.

3.1.7 *sample, n—in methods of chemical analysis,* a portion of a material selected and processed to render its composition representative of the composition of the whole. (Contrast *specimen.*)

3.1.8 *specimen, n—in methods of chemical analysis,* a piece of material selected to be typical of the whole under the assumption that the whole is composed of pieces of similar composition. (Contrast *sample*.)

3.2 *Definitions Referring to Statistics:*

3.2.1 *b-value, n—in statistics* , the difference between the mean of a set of results on a material and its accepted reference value. (Compare *error*.)

3.2.2 *between-laboratory standard deviation,* s_R, *n*—the standard deviation of results obtained on the same material in different laboratories (Synonym: *reproducibility*).

3.2.3 *detection limit, n— for an analytical instrument,* the minimum quantity of analyte expected to yield a response greater than zero.

3.2.4 *error, n—of a result* , the difference between a result obtained on a material and its accepted reference value. (Compare *b-value*.)

3.2.5 *interlaboratory study, ILS, n*—a study undertaken to demonstrate the precision and accuracy of a method.

3.2.6 *minimum standard deviation,* s_M, *n*—the standard deviation of results on a test material obtained under conditions of minimum variability.

3.2.7 *repeatability, n*—see *within-laboratory standard deviation.*

3.2.8 *repeatability standard deviation, n*— see *within-laboratory standard deviation.*

3.2.9 *repeatability index, r, n*—an estimate of the maximum difference expected for results on the same test material on different days in the same laboratory, a difference not expected to be exceeded an average of more than once in 20 comparisons (95 % probability.)

3.2.10 *reproducibility, n*—see *between-laboratory standard deviation.*

3.2.11 *reproducibility standard deviation, n*— see *between-laboratory standard deviation.*

3.2.12 *reproducibility index, R, n*—an estimate of the maximum difference expected for results on the same material in two laboratories, a difference not expected to be exceeded an average of more than once in 20 comparisons (95 % probability.)

3.2.13 *set, n—of results* , a group of results collected under specified conditions for statistical analysis.

3.2.14 *standard deviation, between-laboratory, n*—see *between-laboratory standard deviation.*

3.2.15 *standard deviation, minimum, n*—see *minimum standard deviation.*

3.2.16 *standard deviation, within-laboratory, n*—see *within-laboratory standard deviation.*

3.2.17 *within-laboratory standard deviation,* s_r, *n*—the standard deviation of results collected on the same material in the same laboratory on different days (Synonym: *repeatability*).

3.3 The terms *mean, standard deviation, random* (as in random error), and *systematic error* , in their statistical senses, are adequately defined in Webster's Collegiate Dictionary, Tenth Edition.

4. Analytical Science and Analytical Methods

4.1 Analytical science deals with the development and use of methods for determining chemical composition of materials. Chemical analysis is the application of written analytical methods.

4.2 Analytical method development consists of selecting chemical and physical systems that respond to a specific analyte in a defined suite of material types. The purpose is to define a process that produces a physical change proportional to analyte content

[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~~, Vol 03.05.~~ volume information, refer to the standard's Document Summary page on the ASTM website.

unaffected by other sample components. The measurement system (instrument) yields a numerical result that represents the quantity of analyte. A good analytical method has the following desirable properties:

4.2.1 *Accuracy*—When a method is applied to materials containing various quantities of analyte, it has the property of accuracy if results equal the numerical values of the analyte contents. This property relates solely to a method's average response at each analyte level, ignoring random statistical fluctuations of individual results. Actual methods are never known to be perfectly accurate and this term is usually used in a relative sense to compare different methods or the behavior of a single method under different conditions.

4.2.2 *Precision*—When a method is applied a number of times to a homogeneous sample, it has the property of precision if the result is always the same. This property relates solely to time-related variations in the response of a method and ignores systematic (averaged) differences between results and analyte content that may occur at various analyte levels. Actual methods are never perfectly precise and this term is usually used in a relative sense to compare different methods or the behavior of a single method under different conditions.

4.3 Written methods must satisfy two criteria: (*1*) they shall have the form and editorial style specified in the latest edition of Form and Style for ASTM Standards, and (*2*) the technical content shall be stated in terms that convey precise meanings to laboratory personnel using the method. The language used in the method must direct users (who may not have the same technical knowledge or experience as the developer) to repeat the procedural steps in the manner that produced satisfactory results in the development laboratory. Unless the method conveys this information, users will not achieve the potential accuracy and precision of the method.

4.4 The ILS demonstrates the performance of the method in a group of laboratories typical of those expected to use the method; whereas the accuracy and precision of the analytical techniques and procedures employed are defined by the statistics obtained in the development laboratory. ILS statistics are influenced by three additional factors: (*1*) the success of the translation of the research findings into the method tested in the ILS, (*2*) the care with which the ILS experimental design was followed, and (*3*) the quality of the test materials employed during the ILS.

4.5 The published method contains a summary of the ILS statistics that a user may interpret, based upon the user's own experience. A task group observing a relationship between the method's precision and analyte content as described in Guide E 1763, may choose to provide more detailed descriptions of the method's performance, for example, an equation or table predicting approximate standard deviations at various analyte levels.

5. Statistics and Statistical Methods

5.1 Statistics deals with the collection, analysis, interpretation, and presentation of numerical data sets. The mathematical procedures employed in these processes are statistical methods. Statistic is a generic term for the variable represented by a statistical procedure, but it also refers to the numerical value obtained when a statistical procedure is applied to a data set. For example, mean is a statistic, but the value 6 is the mean of {3, 5, 10}. A statistic is definable by a mathematical expression that calculates to an explicit value for a suitable data set.

5.2 *Statistics Terms*— Refer to common statistics by their common descriptive names.

5.2.1 *Arithmetic average* and *mean* are synonyms.

5.2.2 *Standard deviation* and *precision* are sometimes used indiscriminately. Confusion is avoided if the statistic is always referred to as "the standard deviation." The term "precision" shall be reserved for non-statistical discussions of methods and processes. Use of "precision" when "standard deviation" is meant ~~is~~ to be avoided because these terms have exactly opposite connotations.

5.2.3 *B-value* and *error* are not synonyms. *Total error*, according to common usage, is the net effect of all sources of error. Because error sources must be either "random" or "systematic," total error is the statistical sum of random error and systematic error. For a single result, the error is the difference between the observed value and the accepted (true) value it is intended to estimate. This term requires no qualifying adjective. Lacking additional data, the magnitudes of its component errors are unknown. *B*-value is the difference between the mean of a set of results for a given material and the material's accepted value. It estimates only the systematic error component in the variability of the results in the data set.

5.2.4 A *detection limit* is an instrumental figure of merit used to compare low-level sensitivities of analytical instruments. A number of versions of detection limit have been proposed, all based upon the standard deviation, s_{DL}, of a set of sequential readings on a test material containing little or no analyte. The various detection limits differ primarily in their associated probabilities as defined by a specified multiplier for s_{DL}. It is apparent that s_{DL}, a short-term statistic, cannot include variability caused by calibration operations or by the effects of long-term environmental changes in analytical methods. Thus a detection limit has meaning only for the instrument on which it was determined. It cannot define the performance of methods using the instrument because it is only one of many factors influencing variation in a method's performance in different laboratories.

5.2.5 An *interlaboratory study* (ILS) is a statistically designed demonstration of the actual performance of an analytical method. Practice E 1601 describes the ILS and provides detailed instructions concerning planning and executing the study. The *analysis of variance* (ANOVA) statistical procedure of an ILS provides estimates of:

5.2.5.1 *B-value*—If a certified value for the mean content, *A*, of the analyte in a test material is known, the calculated b-value, B_A, estimates the true difference between the mean and the accepted value at the observed analyte content, $\bar{\bar{x}}$, found in the ILS:

$$B_A = \bar{\bar{x}} - A \quad (1)$$

NOTE 1—How well B_A estimates the theoretically correct inaccuracy of a method is a function of variability in both $\bar{\bar{x}}$ and A ILS on high quality reference materials usually justify use of a simplifying assumption that the variability in A is small enough to be neglected. The variability in $\bar{\bar{x}}$ is normally so large that it yields an illustrative rather than definitive estimate of the true difference.

5.2.5.2 *Minimum Standard Deviation* —For methods involving a completely homogeneous sample (such as a sample solution), minimum variability conditions are defined as replicate results on a sample portion. Under these conditions, the minimum standard deviation, s_M, estimates only instrument performance. For all other methods, minimum variability conditions are defined as results taken on replicate sample portions. Under these conditions, s_M includes variability caused by material inhomogeneity and short-term method variability. In any case, the value of s_M for materials of very low analyte content is an estimate of a pooled s_{DL} of the instruments used in the ILS.

5.2.5.3 *Within-laboratory Standard Deviation and Repeatability Index*—If the task group chooses the ILS design requiring each laboratory to obtain data sets on each of several days, the study produces the within-laboratory standard deviation, s_r, an estimate of the variability from all sources that may operate within a laboratory from day to day. To meet these conditions, the ILS protocol shall ensure that participating laboratories perform every aspect of the method, such as standardization, each day even though they may be required in practice less often. These statistics estimate the variability to be expected in results obtained weeks or months apart. The repeatability index, r, is an estimate, at the 95 % probability level, of the maximum difference for comparing one result with a subsequent result on the same material in the same laboratory.

5.2.5.4 *Between-laboratory Standard Deviation and Reproducibility Index*—The between-laboratory standard deviation, s_R, estimates the variability among determinations made on the same test material in different laboratories. The reproducibility index, R, is an estimate, at the 95 % probability level, of the maximum difference between a result on a material obtained in each of two laboratories.

6. Interpretation of Statistics

6.1 It is beyond the scope of this practice to discuss in detail the subject of development of analytical methods. Statistical methods play key roles in the step-wise processes that culminate in the research version of a method. In the development laboratory, a skilled investigator is able to identify which environmental and method variables need to be controlled and to specify their proper levels to achieve an accurate method. At no other time is a method used under such carefully controlled conditions as during final testing in the development laboratory. The statistics for those results are usually the best basis for evaluating a method's intrinsic accuracy and precision because the investigator takes pains to minimize avoidable sources of variability.

6.2 The published version of a method is produced from the research version in accordance with 4.3-4.5. The ILS is conducted on the publishable method, complete except for the ILS statistics.

6.3 Statistics are properties of the data set from which they derive. A set of results reflects the properties of all the systems and processes that operated to produce it. Statistics summarize a set of results but do not directly describe the behavior of the analytical method that produced them. To make that connection, an investigator attempts to identify the sources of variation in the data and, if possible, to separate the effects that originate in that method from those that do not. A knowledge of statistical principles is helpful in this kind of interpretation, but the task also requires the investigator to be intimately acquainted with analytical science. Normal operations in analytical laboratories, the usual training and experience expected for laboratory personnel, the fundamentals of analytical chemistry, and the details of the analytical method under investigation must all be taken into account.

6.4 The ANOVA model for precision in analytical methods is based upon random variability at three levels: (*1*) The lowest level encompasses short-term factors operating when the method is applied to one sample (a timescale of minutes for most methods). A typical example is the variability of sequentially repeated instrument readings on the same material. (*2*) The intermediate level includes all sources of variability operating within a laboratory. These include low level variability factors plus long-term factors that operate from one day to another in one laboratory. This level includes additional variability from repeated calibration or standardization within a laboratory (a timescale of days, months, or years). (*3*) The highest level of the model includes all variability that may occur between laboratories. This level intrinsically is unrelated to time, but includes, in addition to factors that are time-related, random factors contributing to the observable random average differences among laboratories. Practice E 1601 provides one test design to estimate statistics representative of the lowest and highest levels (Test Plan A) and another providing estimates at all three levels (Test Plan B). If the ILS test protocol adopted by the task group meets the requirements of Practice E 1601 and all participating laboratories follow both the method and the test protocol, the statistics have the meanings and properties described in 5.2.5.

7. Keywords

7.1 performance of analytical methods; statistics in chemical analysis; terms in analytical methods; usage of analytical terms