Designation: E1488 – 23

An American National Standard

Standard Guide for Statistical Procedures to Use in Developing and Applying Test Methods[1]

This standard is issued under the fixed designation E1488; 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 identifies statistical procedures for use in developing new test methods or revising or evaluating existing test methods, or both.

1.2 This guide also cites statistical procedures especially useful in the application of test methods.

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

2. Referenced Documents

2.1 *ASTM Standards:*[2]

E177 Practice for Use of the Terms Precision and Bias in ASTM Test Methods
E178 Practice for Dealing With Outlying Observations
E456 Terminology Relating to Quality and Statistics
E691 Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method
E1169 Practice for Conducting Ruggedness Tests
E1402 Guide for Sampling Design
E2282 Guide for Defining the Test Result of a Test Method
E2489 Practice for Statistical Analysis of One-Sample and Two-Sample Interlaboratory Proficiency Testing Programs
E2554 Practice for Estimating and Monitoring the Uncertainty of Test Results of a Test Method Using Control Chart Techniques
E2586 Practice for Calculating and Using Basic Statistics
E2587 Practice for Use of Control Charts in Statistical Process Control
E2655 Guide for Reporting Uncertainty of Test Results and Use of the Term Measurement Uncertainty in ASTM Test Methods

3. Terminology

3.1 *Definitions*—Unless otherwise noted in this standard, all terms relating to quality and statistics are defined in Terminology E456.

3.1.1 *bias, n*—the difference between the expectation of the test results and an accepted reference value. **E177**

3.1.1.1 *Discussion*—Statistical procedures include the sampling considerations or the experiment design for the collection of data, or both, and the numerical and graphical approaches to summarize and analyze the collected data.

3.1.2 *coefficient of variation, CV, n*—for a nonnegative characteristic, the ratio of the standard deviation to the mean for a population or sample. **E2586**

3.1.3 *component of variance, n*—a part of a total variance identified with a specified source of variability.

3.1.4 *control chart, n*—chart on which are plotted a statistical measure of a subgroup versus time of sampling along with limits based on the statistical distribution of that measure so as to indicate how much common, or chance, cause variation is inherent in the process or product. **E2587**

3.1.5 *observation, n*—the process of obtaining information regarding the presence or absence of an attribute of a test specimen, or of making a reading on a characteristic or dimension of a test specimen. **E2282**

3.1.6 *observed value, n*—the value obtained by making an observation. **E2282**

3.1.7 *precision, n*—the closeness of agreement between independent test results obtained under stipulated conditions. **E177**

3.1.8 *proficiency testing, n*—determination of laboratory testing performance by means of interlaboratory comparisons. **E2489**

3.1.9 *repeatability, n*—precision under repeatability conditions. **E177**

3.1.10 *repeatability conditions, n*—conditions where independent test results are obtained with the same method on

[1] This guide is under the jurisdiction of ASTM Committee E11 on Quality and Statistics and is the direct responsibility of Subcommittee E11.20 on Test Method Evaluation and Quality Control.

Current edition approved Nov. 1, 2023. Published November 2023. Originally approved in 1992. Last previous edition approved in 2023 as E1488 – 12 (2023). DOI: 10.1520/E1488-23.

[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.

identical test items in the same laboratory by the same operator using the same equipment within short intervals of time. **E177**

3.1.11 *repeatability limit r, n*—the value below which the absolute difference between two individual test results obtained under repeatability conditions may be expected to occur with a probability of approximately 0.95 (95 %). **E177**

3.1.12 *repeatability standard deviation, s_r, n*—the standard deviation of test results obtained under repeatability conditions. **E177**

3.1.13 *reproducibility, n*—precision under reproducibility conditions. **E177**

3.1.14 *reproducibility conditions, n*—conditions where test results are obtained with the same method on identical test items in different laboratories with different operators using different equipment. **E177**

3.1.15 *reproducibility limit, R, n*—the value below which the absolute difference between two test results obtained under reproducibility conditions may be expected to occur with a probability of approximately 0.95 (95 %). **E177**

3.1.16 *reproducibility standard deviation, s_R, n*—the standard deviation of test results obtained under reproducibility conditions. **E177**

3.1.17 *ruggedness, n*—insensitivity of a test method to departures from specified test or environmental conditions. **E1169**

3.1.18 *ruggedness test, n*—a planned experiment in which environmental factors or test conditions are deliberately varied in order to evaluate the effects of such variation. **E1169**

3.1.19 *standard deviation, n—of a population*, σ, the square root of the average or expected value of the squared deviation of a variable from its mean – *of a sample* $\bar{x}$, the square root of the sum of the squared deviations of the observed values in the sample divided by the sample size minus 1. **E2586**

3.1.20 *state of statistical control, n*—process condition when only common causes are operating on the process. **E2587**

3.1.21 *statistical procedures, n*—the organized techniques and methods used to collect, analyze, and interpret data.

3.1.21.1 *Discussion*—Statistical procedures include the sampling considerations or the experiment design for the collection of data, or both, and the numerical and graphical approaches to summarize and analyze the collected data.

3.1.22 *test determination, n*—the value of a characteristic or dimension of a single test specimen derived from one or more observed values. **E2282**

3.1.23 *test method, n*—a definitive procedure that produces a test result. **E2282**

3.1.24 *test observation, n*—see **observation**. **E2282**

3.1.25 *test result, n*—the value of a characteristic obtained by carrying out a specified test method. **E2282**

4. Significance and Use

4.1 The creation of a standardized test method generally follows a series of steps from inception to approval and ongoing use. In all such stages there are questions of how well the test method performs.

4.1.1 Assessments of a new or existing test method generally involve statistical planning and analysis. This standard recommends what approaches may be taken and indicates which standards may be used to perform such assessments.

4.2 This standard introduces a series of phases which are recommended to be considered during the life cycle of a test method as depicted in Fig. 1. These begin with a *design phase* where the standard is initially prepared. A *development phase* involves a variety of experiments that allow further refinement and understanding of how the test method performs within a laboratory. In an *evaluation phase* the test method is then examined by way of interlaboratory studies resulting in precision and bias statistics which are published in the standard. Finally, the test method is subject to a *monitoring phase*.

4.3 All ASTM test methods are required to include statements on precision and bias.[3]

4.4 Since ASTM began to require all test methods to have precision and bias statements that are based on interlaboratory studies, there has been increased concern regarding what statistical experiments and procedures to use during the development of the test methods. Although there exists a wide range of statistical procedures, there is a small group of generally accepted techniques that are beneficial to follow. This guide is designed to provide a brief overview of these procedures and to suggest an appropriate sequence of conducting these procedures.

4.5 Statistical procedures often result in interpretations that are not absolutes. Sometimes the information obtained may be inadequate or incomplete, which may lead to additional questions and the need for further experimentation. Information outside the data is also important in establishing standards and in the interpretation of numerical results.

5. Summary of Guide

5.1 Outlined here is a suggested sequence of four phases useful in the development of a test method. A flowchart is provided in Fig. 1. Such a sequence of analyses may require modification in specific situations. The assistance of a qualified statistician is recommended at each review phase.

5.2 *Design Phase:*

5.2.1 This phase includes the formalization of the scope and the significance and use sections. It may include determining the purpose and describing a general approach to the test method but usually does not involve statistical studies.

5.2.2 Studies may be conducted to evaluate the basic performance of the method. The draft test method is prepared and sampling requirements and the test result (see Guide E2282) are clearly defined.

5.2.3 A flow chart is extremely valuable to identify the sequence of operations involved in a test method, for example, the sampling steps required to obtain the test specimens, definition of the test determination, how a test result is to be computed, and running the tests on the specimens.

[3] See the Form and Style Manual for ASTM Standards that specifies, when possible, precision statements shall be estimated based on the results of an interlaboratory test program.

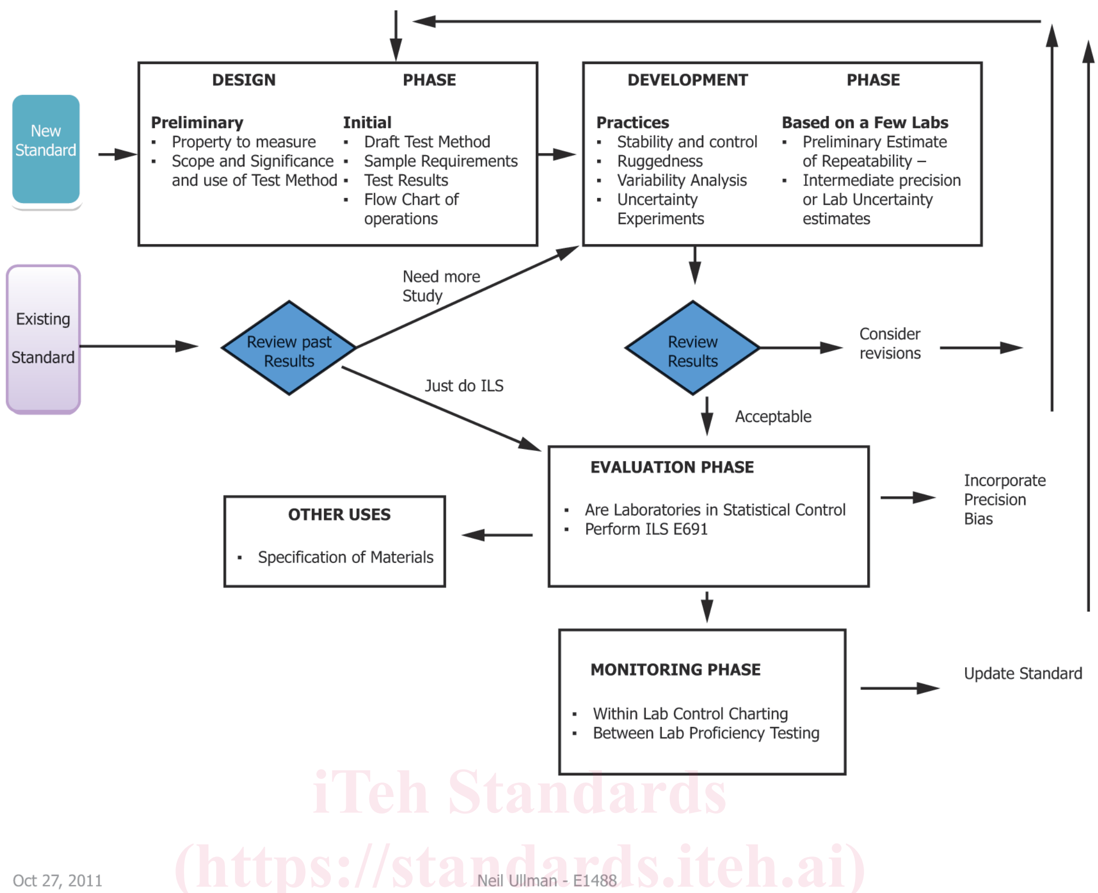

FIG. 1 Sequence of Steps

5.3 *Development Phase:*

5.3.1 The test method is examined for such concerns as its stability, ruggedness, statistical control and the contributions to variability. The completion of this phase should result in preliminary estimates of precision and the identification and suggested ways to estimate potential contributors to uncertainty.

5.3.2 *Evaluation of Short Term Control of Test Method*—A test method must exhibit an ability to provide consistent results at least over short time periods. Preliminary studies or a pilot test should be conducted to evaluate the short term stability of the test method. A small series of repeated tests should be conducted.

5.3.3 *Analysis of Variability*—Statistically designed experiments conducted in one or two laboratories can be used to assess the relative magnitudes of different sources or potential contributors to variability of the test results. Such studies can provide estimates of intermediate measures of precision.

5.3.4 *Ruggedness Test*—A ruggedness test (see Practice E1169) is a statistically designed experiment that helps identify problems in running the test method, clarifies errors, and points out possible environmental conditions, which may adversely affect the test method or point out need for tightening requirements. The ruggedness test can assist in locating ways of reducing variability in the test method.

5.3.5 *Preliminary Estimates of Precision*—From the various studies conducted in accordance with 5.3.2 – 5.3.4, preliminary estimates of repeatability standard deviations should be developed and published in this test method. Until an interlaboratory study is performed, these estimates generally are considered to be provisional. Information on how a laboratory should develop uncertainty estimates should also be provided.

5.3.6 *Statistical Control*—A test method must show capability of performing consistently over time. The use of control charts (see Guide E2655)[3] to monitor a proposed, or existing, test method over time is one recommended way to examine the controllability or stability of a test method. This statistical control should be demonstrated in one or two laboratories using homogeneous material (test specimen).

5.4 *Evaluation Phase:*

5.4.1 The test method is subjected to interlaboratory studies to provide estimates of within-laboratory repeatability and between-laboratory reproducibility. Additional information is supplied from proficiency studies when conducted.

5.4.2 *Interlaboratory Study (ILS)*—In accordance with ASTM Form and Style Manual, whenever feasible, an interlaboratory study must be conducted. This procedure will provide specific estimates of variation anticipated when using the test method.

5.4.3 Protocol for the ILS, Practice E691, provides a guide for developing the ILS for the test method. A first step is the writing of an ILS Protocol, which will define what must be done before the test specimens (or test materials) are distributed to the participating laboratories.

5.4.4 *Precision Statements*—Using the estimates of variation obtained in the ILS, one may prepare precision statements using Practices E691 and E177 or equivalent procedures.

5.5 *Monitoring Phase:*

5.5.1 After a test method is approved and in use it is important to ensure that the published precision and bias statistics for the test method remain achievable and consistent over time or amongst different groups conducting the tests.

5.5.2 *Monitoring Within a Single Location*—It is important for any laboratory or organization that will use a particular test method over time that a means of monitoring to ensure the method results using quality control samples are stable and in control. Regular evaluation of the uncertainty (Practice E2554) or use of a control charting method (Practice E2587) are two ways to monitor the test method.

5.5.3 *Between Laboratory Comparisons*—Proficiency testing programs measure the typical variation amongst various laboratories. The specific laboratories involved also obtain information about how well they perform compared to other laboratories.

6. Development of Test Method

6.1 Proposed standards that are under development should be treated in a formal manner following as many of the suggested procedures as possible. Standards that are already in existence as approved test methods or in general practice require periodic review that would include selected procedures.

6.2 *Under Development*—The development stage involves test methods that are in the preliminary stages during which equipment may not have been fully tested, practices are not agreed upon, and operators have yet to be adequately trained. Often this stage also applies to standards that have not yet been approved.

6.2.1 It is essential that tests for statistical control, ruggedness, and variability analyses be conducted prior to any interlaboratory test programs.

6.2.2 After all major environmental contributors have been identified, controlled, and incorporated into the test method, and after adequate standardized equipment is available, an ILS can be conducted. The ILS must be completed prior to the first five-year review. The committee should strive to have ILS results as soon as possible.

6.2.3 After evaluating data from ruggedness tests, variability analysis, or an ILS, changes to the test method may be suggested.

6.2.4 If major changes are made to the test method, a repeat of the various steps is usually necessary. Precision and bias statements should reflect the most current version of the test method.

6.3 *Existing Standards*—These standards comprise test methods that are in common use for which standard equipment may exist and for which experienced operators have been trained and are available.

6.3.1 Control charting, ruggedness tests, and variability analyses will be useful, especially if they have not previously been conducted. Such tests may provide better information about variation and necessary tolerances than has previously been available.

6.3.2 If precision estimates have not been established through an actual interlaboratory test program, then such a program should be initiated.

7. Data and Sampling

7.1 *Sample Determination:*

7.1.1 The sampling section of a standard should indicate clearly what constitutes the primary sampling unit, how that sampling unit is further subdivided, and how multiple test values are designated. (See Guide E1402.)

7.1.2 In considering the implication of test results as they relate to the material, the test method should be clear whether the sampling method or the test is destructive or nondestructive.

7.1.3 The user of the test method should be aware of whether the standard calls for a random sample. In some standards, as for example in sampling from coils or rolls of material, samples may be taken only from certain portions of the material.

7.2 *Test Result Determination*—The procedure for determining a test result must be clear and unambiguous.

7.2.1 An observation leads to an observed value.

7.2.2 Several observed values may lead to a test determination. The observed values need not be the same type of measurements (for example, they may consist of three readings such as length, width, and mass).

7.2.3 Several test determinations may lead to a test result, as by averaging three test determinations.

7.2.4 A test result is the consequence of a single execution of the entire test method.

7.3 *Type of Data*—The kind of data that results from the application of the test method determines the types of statistical analyses to be performed.

7.3.1 *Numerical versus Categorical/Attribute Data*—Most of the statistical procedures referred to in this guide deal with numerical data. Control charts are available for all types of data, but all ILS procedures currently in use depend on numerical data.

7.3.2 *"Normally" Distributed Data*—Most of the statistical procedures referred to in this guide consider that the unknown distribution of the test results can be modeled by a normal distribution.

8. Sources of Variability

8.1 *Experimental Realization of a Test Method:*

8.1.1 A realization of a test method refers to an actual application of the test method to produce a test result as specified by the test method. The realization involves an interpretation of the written document by an operator, who uses a specific unit and version of the specified test apparatus, in the particular environment of a testing laboratory, to evaluate a specified number of test specimens of the material to be tested. Another realization of the test method may involve a change in one or more of the experimental factors listed above. The test result obtained by another realization of the test method will