Designation: E2027 – 09

Standard Practice for Conducting Proficiency Tests in the Chemical Analysis of Metals, Ores, and Related Materials[1]

This standard is issued under the fixed designation E2027; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This practice provides direction for organizing and conducting proficiency test programs in analytical chemistry for metals, ores, and related materials. It is consistent with ISO Guide 43 and Guide E1301. It does not address the selection and use of proficiency testing schemes by accrediting bodies.

2. Referenced Documents

2.1 *ASTM Standards:*[2]

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

E826 Practice for Testing Homogeneity of a Metal Lot or Batch in Solid Form by Spark Atomic Emission Spectrometry

E1187 Terminology Relating to Conformity Assessment (Withdrawn 2006)[3]

E1301 Guide for Proficiency Testing by Interlaboratory Comparisons (Withdrawn 2012)[3]

E1724 Guide for Testing and Certification of Metal, Ore, and Metal-Related Reference Materials (Withdrawn 2010)[3]

2.2 *ISO Standards:*[4]

ISO 17025 General Requirements for the Competence of Calibration and Testing Laboratories

ISO Guide 43 Proficiency Testing by Interlaboratory Comparisons

ISO Guide 9000 Quality Management and Quality System Elements

3. Terminology

3.1 *Definitions*—For formal definitions related to laboratory accreditation, Terminology E1187 applies.

3.2 For other definitions of terms used in the practice, refer to Terminology E135.

4. Significance and Use

4.1 This practice sets the basic requirements for proficiency test programs in the chemical analysis of metals, ores, and related materials. It does not set specific procedural requirements, but does establish a framework for particular programs, including those with either small or large numbers of participants. (**Warning—**The data from proficiency testing programs must never be used to assign certification values to the materials used in the program. The elements of a properly conceived and implemented certification program are described in detail in Guide E1724.)

4.2 Most accreditation bodies require that laboratories participate regularly in proficiency testing programs that they have accepted for the purpose. Therefore, it is essential that each program comply with accepted principles including technical requirements, statistical procedures (see Annex A1), and quality management (see Annex A2).

5. Types of Proficiency Testing

5.1 Proficiency testing techniques vary depending on the nature of the test item, the method in use and the number of laboratories participating. The most common approach involves randomly selected sub-samples from a source of material being distributed simultaneously to participating testing laboratories for concurrent testing. It is essential that all of the material from which the participants' test materials are taken be sufficiently homogeneous so that any results later identified as outliers should not be attributed to any significant test item variability. After completion of the testing, the results are returned to the coordinating body, and compared with the assigned value(s) or to the mean and standard deviations obtained from a statistical analysis of the data to give an indication of the performance of the individual laboratories and the group as a whole.

5.2 In some cases, separate portions of previously certified reference materials are circulated.

[1] This practice 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, 2009. Published December 2009. Originally approved in 1999. Last previous edition approved in 2005 as E2027 – 05. DOI: 10.1520/E2027-09.

[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.

[3] The last approved version of this historical standard is referenced on www.astm.org.

[4] Available from American National Standards Institute (ANSI), 25 W. 43rd St., 4th Floor, New York, NY 10036, http://www.ansi.org.

6. Organization and Design

6.1 *Framework:*

6.1.1 The design stage of any proficiency testing program requires the input of technical experts, statisticians and a program coordinator to ensure its success and smooth operation.

6.1.2 The coordinator, in consultation with these other personnel, develops a program appropriate to the particular proficiency test. A proficiency test program shall be designed to avoid any confusion about its objectives. A plan shall be established and documented (see Annex A2) before the start of the program and shall include the following information:

6.1.2.1 The name and the address of the organization conducting the proficiency program,

6.1.2.2 The name and address of the coordinator and other personnel involved in the design and operation of the proficiency program,

6.1.2.3 The nature and the purpose of the proficiency program,

6.1.2.4 A procedure for the manner in which the participants are selected, or criteria that need to be met before participation is allowed,

6.1.2.5 The name and address of the laboratory or laboratories performing the various parts of the program (for example, sampling, sample processing, homogeneity testing and assigning values) and a description of the market to be served,

6.1.2.6 The nature of the test material(s) and test(s) selected, as well as a short description of the considerations underlying these choices,

6.1.2.7 A description of the manner in which the test materials are obtained, processed, checked and transported,

6.1.2.8 The time schedule for the various phases of the proficiency testing,

6.1.2.9 The expected initial and target dates or deadlines of the proficiency program including the date(s) for the testing to be conducted by the participants,

6.1.2.10 For ongoing programs, the frequency at which test materials are distributed,

6.1.2.11 Information on methods or procedures which participants may need to use to perform the tests or measurements (ASTM test methods, laboratory standard procedures/methods, etc.),

6.1.2.12 An outline of the statistical analysis to be used including the determination of assigned value(s) and any outlier detection techniques,

6.1.2.13 The basis for performance evaluation techniques, and

6.1.2.14 A description of the extent to which the test results, and the conclusions that will be based on the outcome of the proficiency tests, are to be made public.

6.2 *Staff:*

6.2.1 The staff shall include, or collaborate closely with, those holding adequate qualifications and experience in the design, implementation and reporting of interlaboratory comparisons. They shall possess appropriate technical, statistical and administrative skills.

6.2.2 The operation of specific interlaboratory comparisons requires the guidance of persons with detailed technical knowledge and experience of the test methods involved. To this end the coordinator shall enlist some professionals to act as an advisory group. The functions of this advisory group may be to:

6.2.2.1 Develop and review procedures for the planning execution, analysis, reporting and monitoring the effectiveness of the proficiency testing program,

6.2.2.2 Identify and evaluate interlaboratory comparisons organized by other bodies,

6.2.2.3 Evaluate proficiency test results of participating laboratories,

6.2.2.4 Provide advice to any body assessing the technical competence of participating laboratories, both on the results obtained during a proficiency test program, and how those results should be used with other aspects of laboratory evaluations,

6.2.2.5 Provide advice to participants who appear to experience problems, and

6.2.2.6 Resolve disputes between the coordinator and participants.

6.3 *Data Processing Equipment*—Equipment shall be adequate to conduct all necessary data entry and statistical analyses and provide timely and valid results. Procedures for checking data entry shall be implemented and all software shall be verified, supported and backed up. The storage and security of data files shall be controlled.

6.4 *Statistical Design:*

6.4.1 The statistical model and data analysis techniques to be used shall be documented together with a short description of the background to their selection. Further details of common statistical procedures and treatment of proficiency testing data are discussed in Annex A1.

6.4.2 Careful consideration shall be given to the following matters and their interactions: the repeatability and reproducibility of the test(s) involved; the smallest differences to be detected between participating laboratories at a desired confidence level; the number of participating laboratories; the number of samples to be tested and the number of repeat tests or measurements to be carried out on each sample; the procedures to be used to estimate the assigned value; procedures to be used to identify outliers; and, potential bias in the test methods employed.

6.5 *Test Materials Preparation:*

6.5.1 Preparation of test materials may either be outsourced or undertaken by the coordinator. The organization preparing the test material shall have demonstrable competence to do so.

6.5.2 Any conditions relating to the test materials that may affect the integrity of the interlaboratory comparison, such as homogeneity, stability, possible damage in transit and effects of ambient conditions shall be considered.

6.5.3 The test materials or materials to be distributed in the program shall be similar in nature to those routinely tested by participating laboratories.

6.5.4 The number of test materials to be distributed may depend on whether there is a requirement to cover a range of compositions.

6.5.5 The assigned value(s) shall not be disclosed to the participants until after the results have been collated. However, in some cases it may be appropriate to advise target ranges prior to testing.

6.5.6 Consideration may be given to preparation of additional test materials other than those needed for the proficiency test program. Surplus test materials may be useful as quality control materials, test samples for interlaboratory tests of new test methods, or training aids for laboratories after results from participants have been evaluated.

6.6 *Test Materials Management:*

6.6.1 Procedures for sampling, randomizing, transporting, receiving, identifying, labelling, storing and handling of test materials shall be documented.

6.6.2 Where bulk material is prepared for a proficiency test, it shall be sufficiently homogeneous for each test parameter so that all laboratories will receive test materials that do not differ significantly in the parameters to be measured. The coordinator shall clearly state the procedure used to establish the homogeneity of the test item (see A1.4). Homogeneity testing shall be conducted prior to the dispatch of the test materials to the participating laboratories.

6.6.3 Where applicable, the coordinating laboratory shall also provide evidence that the test materials are sufficiently stable to ensure that they will not undergo any significant change throughout the conduct of the proficiency test. When unstable analytes need to be assessed, it may be necessary for the coordinating organization to prescribe a date by which the testing shall be completed, including required special pretesting procedures.

6.6.4 Coordinators shall consider any hazards that the test materials might pose and take appropriate action to advise any party that might be at risk (for example, test material distributors, testing laboratories, etc.).

6.7 *Choice of Test Method:*

6.7.1 The coordinator may instruct participants to use a specified test method. Such test methods are usually nationally or internationally-accepted standard test methods, and will have been validated by an appropriate procedure (for example, collaborative trial).

6.7.2 Participants may be able to use the test method of their choice, which is consistent with routine procedures used in their laboratories. Where participants are free to use a test method of their own choice, coordinators shall request details of the test methods used to allow, where appropriate, the use of participants' results to compare and comment on the test methods.

7. Operation and Reporting

7.1 *Coordination and Documentation*—The day-to-day operation of a program shall be the responsibility of a coordinator. All practices and procedures shall be documented. These may be incorporated in, or supplemented by, a quality manual (see Annex A2).

7.2 *Instructions:*

7.2.1 Detailed instructions covering all aspects of the program that should be followed by the participating laboratories shall be provided. These may be provided, for example, as an integral part of a program protocol.

7.2.2 Instructions shall include details concerning factors that could influence the testing of the supplied materials. Such factors shall include qualifications of operators, nature of the materials, equipment status, selection of test procedures and timing of testing.

7.2.3 Specific instructions on the recording and reporting of test or calibration results shall also be supplied (for example, units, number of significant figures, reporting basis, result deadlines, etc.).

7.2.4 Participants shall be advised to treat proficiency testing items as if they were routine tests (unless there are some special requirements in the design of the proficiency test which may require departure from this principle). They shall also be advised to avoid collusion with other participants.

7.2.5 Participants shall be advised to ensure that their laboratory capabilities are compatible with the protocols and test samples provided by the programs. Incompatibility between the program and its participants' capabilities can lead to inappropriate indicators of poor performance.

7.3 *Packaging and Transportation:*

7.3.1 The coordinator of the program shall ensure that packaging and methods of transport are adequate and able to protect the stability and characteristics of the test materials. There may be certain restrictions on transportation such as dangerous goods regulations, or customs requirements. In some cases, the laboratories themselves also take responsibility for the transport of the items, particularly in sequential measurement comparisons programs.

7.3.2 All appropriate customs declaration forms shall be completed by the coordinator to ensure that delays in customs clearance are minimized. The program shall comply with national and international regulations applicable to test item transport.

7.4 *Data Analysis and Records:*

7.4.1 The results received from the participating laboratories shall be entered and analyzed and then reported as soon as practicable. It is essential that procedures are implemented to check the validity of data entry and transfers and subsequent statistical analysis. Data sheets, computer back-up files, printouts, graphs, etc., shall be retained for a specified period.

7.4.2 Data analysis shall generate summary measures and performance statistics and associated information consistent with the program's statistical model and objectives. The influence of extreme results on summary statistics shall be minimized by the use of outlier detection tests to identify and then omit them or, preferably, by the use of robust statistics. Annex A1 contains some broad suggestions for statistical evaluations.

7.4.3 Program coordinators shall have documented criteria for dealing with test results that may be inappropriate for proficiency evaluations. For example, it is recommended that for analytes for which the test material has been shown not to be sufficiently homogeneous or stable for the purposes of a proficiency test, no grading or scoring shall be given for those analytes.

7.5 *Program Reports:*

7.5.1 The content of program reports will vary depending on the purpose of a particular program, but shall be clear and comprehensive and include data on the distribution of results from all laboratories together with an indication of individual participant's performance.

7.5.2 The following information shall be included in reports of proficiency programs:

7.5.2.1 Name and address of the organization conducting or coordinating the program,

7.5.2.2 Names and affiliations of persons involved in the design and conduct of the program,

7.5.2.3 Date of issue of report,

7.5.2.4 Report number and clear identification of program,

7.5.2.5 Clear description of items or materials used including details of sample preparation and homogeneity testing,

7.5.2.6 Laboratory participation codes and test results,

7.5.2.7 Statistical data and summaries including assigned values and range of acceptable results,

7.5.2.8 Procedures used to establish any assigned value,

7.5.2.9 Details of the traceability and uncertainty of any assigned value,

7.5.2.10 Assigned values and summary statistics for test methods or procedures used by other participating laboratories (if different test methods are used by different laboratories),

7.5.2.11 Comments on laboratory performance by the coordinator and technical advisers,

7.5.2.12 Procedures used to design and implement the program (which may include reference to a program protocol),

7.5.2.13 Procedures used to statistically analyze the data (see Annex A1), and

7.5.2.14 Advice, where appropriate, on the interpretation of the statistical analysis.

7.5.3 For programs operated on a regular basis, it may be sufficient to have simpler reports such that many of the recommended elements in 7.5.2 could be excluded from routine reports, but included in periodic summary reports and on request from participants.

7.5.4 Reports shall be made available quickly within specified timetables. All original data supplied shall be reported to participants. In some programs, such as long period measurement comparison programs, interim reports shall be issued to individual participants.

7.6 *Evaluation of Performance:*

7.6.1 The coordinator shall retain control over the evaluation of performance to ensure the credibility of the program.

7.6.2 The coordinator shall enlist the assistance of technical advisers to provide expert commentary on performance with respect to:

7.6.2.1 Overall performance versus prior expectations (taking uncertainties into account),

7.6.2.2 Variation within and between laboratories (and comparisons with any previous programs or published precision data),

7.6.2.3 Variation between test methods or procedures, if applicable,

7.6.2.4 Possible sources of error and suggestions for improving performance,

7.6.2.5 Any other suggestions, recommendations or general comments, and,

7.6.2.6 Conclusions.

7.6.3 It may be helpful to provide individual summary sheets for participants periodically during or after a particular program and these may include updated summaries of performance of individual laboratories over various rounds of an ongoing program. Such summaries can be further analyzed and trends highlighted, if required.

7.6.4 A variety of procedures exist to assess performance of participants. Some examples of procedures are given in Annex A1.

7.6.5 Reporting of performance by ranking laboratories in a table according to their performance is not recommended in proficiency testing. Therefore, ranking shall only be used with extreme caution, as it can be misleading and open to misinterpretation.

7.7 *Communication with Participants:*

7.7.1 Participants shall be provided with a detailed set of information upon joining a proficiency testing program, such as a formal program protocol. Subsequent communication with participants may be by letter, newsletter or reports, or a combination thereof, together with periodic meetings. Participants shall be advised immediately of any changes in program design or operation.

7.7.2 Participants shall be able to contact the coordinator if they consider that assessment of their performance in a proficiency test is in error.

7.7.3 Feedback from laboratories shall be encouraged, so that participants actively contribute to the development of a program.

8. Confidentiality and Ethical Considerations

8.1 *Confidentiality of Records:*

8.1.1 Programs shall maintain confidentiality of the identity of the data associated with individual participants. In some circumstances, a coordinating body may be required to report poor performance to a particular authority, but participants shall be notified of this possibility.

8.1.2 A group of participants may elect to waive confidentiality within the group, for the purposes of discussion and mutual assistance in improvement.

8.2 *Collusion and Falsification of Results:*

8.2.1 Although proficiency testing programs are intended primarily to help participants improve their performance, there may be a tendency among some participants to provide a falsely optimistic impression of their capabilities. For example, collusion may take place between laboratories, so that truly independent data are not submitted. Laboratories may also give a false impression of their performance if they routinely carry out single analyses, but report the mean of replicate determinations on the proficiency test materials or conduct additional replicates to those specified for a particular program. Proficiency testing programs shall be designed to minimize collusion and falsification.

8.2.2 Although all reasonable measures shall be taken by the coordinators to prevent collusion, it shall be understood that it is the responsibility of the participating laboratories to avoid it.

9. Keywords

9.1 practice; proficiency testing; protocol

ANNEXES

(Mandatory Information)

A1. STATISTICAL METHODS FOR TREATMENT OF PROFICIENCY TEST DATA

INTRODUCTION

Proficiency test results can appear in many forms, spanning a wide range of data types and underlying statistical distributions. The statistical techniques used to analyze the results need to be appropriate for each situation, and so are too varied to prescribe. There are, however, three steps common to all proficiency tests, when participants' results are to be evaluated: determination of the assigned value, calculation of performance statistics, evaluation of performance, and, in some cases, preliminary determination of test material homogeneity.

Annex A1 gives general criteria for statistical techniques that can be applied as needed to guide specific applications.

With new interlaboratory comparison programs, agreement is often poor due to new questions, new forms, artificial test materials, poor agreement of test methods, or variable laboratory procedures. Coordinators may have to use robust measures of relative performance (such as percentiles) until agreement improves. Statistical techniques may need to be refined once interlaboratory agreement has improved and proficiency testing is well established.

A1.1 Determination of the Assigned Value and Its Uncertainty

A1.1.1 There are various procedures available for the establishment of assigned values. The most common procedures are listed below in an order that, in most cases, will result in increasing uncertainty for the assigned value:

A1.1.1.1 *Known Values*—With results determined by specific test material formulation (for example, manufacture or dilution).

A1.1.1.2 *Certified Reference Values*—As determined by definitive methods (for quantitative tests).

A1.1.1.3 *Reference Values*—As determined by analysis, measurement or comparison of the test material alongside a reference material or standard, traceable to a national or international standard.

A1.1.1.4 *Consensus Values from Expert Laboratories*—Expert laboratories shall have demonstrable competence in the determination of the analytes under test using validated test methods known to be highly precise and accurate, and comparable to test methods in general use.

A1.1.1.5 *Consensus Values from Participant Laboratories.*

A1.1.2 Assigned values shall be determined to evaluate participants fairly, yet to encourage interlaboratory and intermethod agreement. This is accomplished through selection of common comparison groups wherever possible, and the use of common assigned values.

A1.1.3 The following statistics may be appropriate when assigned values are determined by consensus techniques:

A1.1.3.1 *Qualitative Value*—Consensus of a predetermined majority percentage (usually expressed on a nominal scale).

A1.1.3.2 *Quantitative Value*—"Average" for an appropriate comparison group such as mean, which may be weighted or transformed (for example, trimmed or geometric mean), or median, mode or other robust measure.

A1.1.4 *Extreme Results:*

A1.1.4.1 When participants' results are used to determine assigned values, techniques should be in place to minimize the influence of extreme results. This can be accomplished with robust statistical methods or by removing outliers prior to calculation. In larger or routine programs, it may be possible to have automated outlier screens.

A1.1.4.2 If results are removed as outliers, they shall be removed only for calculation of summary statistics. These results shall still be evaluated within the proficiency program but be given the lowest performance rating.

A1.1.5 *Other Considerations:*

A1.1.5.1 Ideally, if assigned values are determined by reference or participant consensus, the coordinator shall have a procedure to establish the trueness of the assigned values, and,

A1.1.5.2 The coordinator shall have criteria for the acceptability of an assigned value in terms of its uncertainty.

A1.2 Calculation of Performance Statistics

PERFORMANCE ON SINGLE TEST MATERIALS

A1.2.1 Proficiency test results often need to be transformed into a performance statistic to aid interpretation and to allow comparison with defined goals. The objective is to measure the deviation from the assigned value in a manner that allows

comparison with performance criteria. Techniques may range from no processing required to complex statistical transformations.

A1.2.2 Performance measures shall be meaningful to program participants. Therefore, measures shall relate to the application needs for the test and be well understood or traditional within a particular field.

A1.2.3 Variability measures are often used for calculation of performance statistics and in summary reports of proficiency testing programs. Common examples of such variability measures for an appropriate comparison group include:

A1.2.3.1 Standard deviation (Std Dev),

A1.2.3.2 Relative standard deviation $\left(RSD = \frac{\text{Std Dev}}{\text{mean}} \times 100\right)$, and

A1.2.3.3 Percentiles, median absolute deviation or other robust measure.

A1.2.4 For qualitative results, no calculation is usually necessary. Commonly used statistics for quantitative results are listed below in increasing degree of transformation of participants' results:

A1.2.4.1 Difference, $(x–X)$, where x is participant's result and X is the assigned value,

A1.2.4.2 Percent difference, $\frac{(x-X)}{X} \times 100$,

A1.2.4.3 Percentile or rank,

A1.2.4.4 Transformed difference, such as $(x–X)^2$ or log $(1 + | x–X |)$, and

A1.2.4.5 z scores, where $z = \frac{x-X}{s}$, and s is an estimate or measure of variability. This model can be used both in the situation where X and s are derived from participants' results or when X and s are not derived from all the participant results.

A1.2.5 *Considerations:*

A1.2.5.1 The simple difference between the participant result and the assigned value may be adequate to determine performance, and is most easily understood by participants.

A1.2.5.2 The percent difference adjusts for concentration, and is well understood by participants.

A1.2.5.3 Percentiles or ranks are useful for highly disperse or skewed results, ordinal responses, or when there are a limited number of different responses. This technique should be used with caution.

A1.2.5.4 Transformed results may be preferred, or necessary, depending on the nature of the test. For example, dilution–based results are a form of geometric scale, transformable by logarithms.

A1.2.5.5 If statistical criteria are used (for example, z scores), the estimates of variability shall be reliable, that is, based on enough observations to reduce the influence of extreme results and achieve low uncertainty.

COMBINED PERFORMANCE SCORES

A1.2.6 Performance may be evaluated on the basis of more than one result in a single proficiency test round. This might occur when there is more than one test material for a particular analyte, or a family of related analytes. This would be done to provide a more comprehensive evaluation of performance. Examples are:

A1.2.6.1 Composite score for the same analyte: number of satisfactory results, average absolute z score, average absolute difference (in units or percent), summed absolute difference (or squared difference), rescaled sum of z scores (RSZ), rescaled sum of squared z scores (RSSZ), and precision.

A1.2.6.2 Composite score for different analytes: number (or percent) of satisfactory results, and average absolute z score.

A1.2.7 *Considerations:*

A1.2.7.1 Scores may be transformed (if necessary) so that they all follow the same assumed distribution (for example, Gaussian for z scores, or Chi Square for squared differences).

A1.2.7.2 There shall be a check for extreme values that could heavily influence a quantitative composite score.

A1.3 Evaluation of Performance

INITIAL PERFORMANCE

A1.3.1 Criteria for performance evaluation should be established after taking into account whether the performance measure involves the following features:

A1.3.1.1 *Expert Consensus*—Where the advisory group, or other qualified experts directly determine whether reported results are fit for purpose. Expert consensus is the typical way to assess results for qualitative tests.

A1.3.1.2 *Fitness for Purpose*—Considering, for example, test method performance specifications and participants' recognized level of operation.

A1.3.1.3 *Statistical Determination for Scores*—Where criteria should be appropriate for each score, for example, for z scores:

$|z| \leq 2$ = satisfactory
$2 < |z| < 3$ = questionable
$|z| \geq 3$ = unsatisfactory

A1.3.1.4 *Consensus of Participants*—The range of scores or results used by some percentage of participants, or from a reference group, such as: central percentage (80, 90, or 95 %) satisfactory, or one-sided percentage (lowest 90 %) satisfactory.

A1.3.2 For split sample designs, an objective may be to identify inadequate calibration or large random fluctuation, or both, in results. In these cases, evaluations should be based on an adequate number of results and across a wide range of concentrations. Graphical techniques are useful for identifying and describing these problems. Results can be compared using regression analysis, with appropriate parametric or nonparametric techniques.

A1.3.3 Graphs should be used whenever possible to show performance (for example, histograms, error bar charts, ordered z score charts). These charts can be used to show distributions of participant values, relationship between results on multiple test materials, and comparative distributions for different test methods.

MONITORING PERFORMANCE OVER TIME

A1.3.4 A proficiency test program can include techniques to monitor performance over time. The statistical techniques should allow participants to see the variability in their performance; whether there are general trends or inconsistencies, and where the performance varies randomly.

A1.3.5 Graphical methods should be used to facilitate interpretation by a wider variety of readers. Traditional Shewhart control charts are useful, particularly for self-improvement purposes. Data listings and summary statistics allow more detailed review. Statistics used to evaluate performance should be used for these graphs and tables.

A1.4 *Preliminary Determination of Test Material Homogeneity*—Appropriate statistical techniques shall be used for the evaluation of data from homogeneity testing of test materials. One suitable approach for metals testing is Practice E826.

A2. QUALITY ASSURANCE OF PROFICIENCY TESTING PROGRAMS

A2.1 A quality assurance system shall be established and maintained. This system shall be documented, for example, in a quality manual. It shall outline the policies and procedures which exist to ensure the quality of the proficiency testing programs, to give confidence to both participants and users of participants' data. It is recommended that the organization conducting a program shall meet the requirements of quality assurance and technical competence based on the appropriate parts of the ISO Guide 9000 series and ISO 17025 as demonstrated by certification or accreditation, or both, by a recognized body, where available.

A2.1.1 The following topics shall be included in such quality documentation:

A2.1.1.1 Quality policy,

A2.1.1.2 Organization of coordinating body,

A2.1.1.3 Staff roles and responsibilities,

A2.1.1.4 Documentation control,

A2.1.1.5 Audit and review procedures,

A2.1.1.6 Aims, scope, statistical design and format of proficiency testing programs,

A2.1.1.7 Operational procedures, including: sample preparation; homogeneity testing of samples; equipment; procedures for establishing assigned values; suppliers (including subcontractors); logistics; and analysis of data,

A2.1.1.8 Preparation and issuing of reports,

A2.1.1.9 Action and feedback by participants,

A2.1.1.10 Documentation of records,

A2.1.1.11 Complaint handling procedures,

A2.1.1.12 Policies on confidentiality and ethical procedures,

A2.1.1.13 Computing information,

A2.1.1.14 Safety and other environmental factors,

A2.1.1.15 Sub-contracting,

A2.1.1.16 Fees for participation,

A2.1.1.17 Scope of availability of programs, and

A2.1.1.18 General policies on participation and on use of program results.