EP15-A2
Vol. 25 No. 17
Replaces EP15-A
Vol. 21 No. 25

User Verification of Performance for Precision and Trueness; Approved Guideline—Second Edition

This document describes the demonstration of method precision and trueness for clinical laboratory quantitative methods utilizing a protocol designed to be completed within five working days or less.

A guideline for global application developed through the Clinical and Laboratory Standards Institute consensus process.

CLINICAL AND LABORATORY STANDARDS INSTITUTE ™

(Formerly NCCLS)
Providing NCCLS standards and guidelines,
ISO/TC 212 standards, and ISO/TC 76 standards

Clinical and Laboratory Standards Institute

Providing NCCLS standards and guidelines, ISO/TC 212 standards, and ISO/TC 76 standards

The Clinical and Laboratory Standards Institute (CLSI, formerly NCCLS) is an international, interdisciplinary, nonprofit, standards-developing, and educational organization that promotes the development and use of voluntary consensus standards and guidelines within the healthcare community. It is recognized worldwide for the application of its unique consensus process in the development of standards and guidelines for patient testing and related healthcare issues. Our process is based on the principle that consensus is an effective and cost-effective way to improve patient testing and healthcare services.

In addition to developing and promoting the use of voluntary consensus standards and guidelines, we provide an open and unbiased forum to address critical issues affecting the quality of patient testing and health care.

PUBLICATIONS

A document is published as a standard, guideline, or committee report.

Standard A document developed through the consensus process that clearly identifies specific, essential requirements for materials, methods, or practices for use in an unmodified form. A standard may, in addition, contain discretionary elements, which are clearly identified.

Guideline A document developed through the consensus process describing criteria for a general operating practice, procedure, or material for voluntary use. A guideline may be used as written or modified by the user to fit specific needs.

Report A document that has not been subjected to consensus review and is released by the Board of Directors.

CONSENSUS PROCESS

The CLSI voluntary consensus process is a protocol establishing formal criteria for:

- the authorization of a project
- the development and open review of documents
- the revision of documents in response to comments by users
- the acceptance of a document as a consensus standard or guideline.

Most documents are subject to two levels of consensus—"proposed" and "approved." Depending on the need for field evaluation or data collection, documents may also be made available for review at an intermediate consensus level.

Proposed A consensus document undergoes the first stage of review by the healthcare community as a proposed standard or guideline. The document should receive a wide and thorough technical review, including an overall review of its scope, approach, and utility, and a line-by-line review of its technical and editorial content.

Approved An approved standard or guideline has achieved consensus within the healthcare community. It should be reviewed to assess the utility of the final document, to ensure attainment of consensus (i.e., that comments on earlier versions have been satisfactorily addressed), and to identify the need for additional consensus documents.

Our standards and guidelines represent a consensus opinion on good practices and reflect the substantial agreement by materially affected, competent, and interested parties obtained by following CLSI's established consensus procedures. Provisions in CLSI standards and guidelines may be more or less stringent than applicable regulations. Consequently, conformance to this voluntary consensus document does not relieve the user of responsibility for compliance with applicable regulations.

COMMENTS

The comments of users are essential to the consensus process. Anyone may submit a comment, and all comments are addressed, according to the consensus process, by the committee that wrote the document. All comments, including those that result in a change to the document when published at the next consensus level and those that do not result in a change, are responded to by the committee in an appendix to the document. Readers are strongly encouraged to comment in any form and at any time on any document. Address comments to the Clinical and Laboratory Standards Institute, 940 West Valley Road, Suite 1400, Wayne, PA 19087, USA.

VOLUNTEER PARTICIPATION

Healthcare professionals in all specialties are urged to volunteer for participation in CLSI projects. Please contact us at customerservice@clsi.org or +610.688.0100 for additional information on committee participation.

EP15-A2
ISBN 1-56238-574-7
Volume 25 Number 17 ISSN 0273-3099

User Verification of Performance for Precision and Trueness; Approved Guideline—Second Edition

R. Neill Carey, PhD
F. Philip Anderson, PhD
Harvey George, PhD
Alfred E. Hartmann, MD
Verlin K. Janzen, MD
Anders Kallner, MD, PhD
Jacob B. Levine, MBA
Jennifer Schiffgens, MBA, MT(ASCP)
Aparna Srinivasan, PhD
Daniel W. Tholen, MS

Abstract

Clinical and Laboratory Standards Institute document EP15-A2—*User Verification of Performance for Precision and Trueness; Approved Guideline—Second Edition* describes the demonstration of method precision and trueness for quantitative methods performed within the laboratory. Included are guidelines for the duration, procedures, materials, data summaries, and interpretation techniques that are adaptable for the widest possible range of analytes and device complexity. A balance is created in the document between the complexity of design and formulae, and the simplicity of operation. The protocol is designed to be completed within five working days. Definitions are provided for repeatability, within-laboratory precision, and other terms and concepts used in the document.

Clinical and Laboratory Standards Institute (CLSI). *User Verification of Performance for Precision and Trueness; Approved Guideline—Second Edition*. CLSI document EP15-A2 (ISBN 1-56238-574-7). Clinical and Laboratory Standards Institute, 940 West Valley Road, Suite 1400, Wayne, Pennsylvania 19087-1898 USA, 2005.

(Formerly NCCLS)
Providing NCCLS standards and guidelines,
ISO/TC 212 standards, and ISO/TC 76 standards

Suggested Citation

(Clinical and Laboratory Standards Institute. *User Verification of Performance for Precision and Trueness; Approved Guideline—Second Edition.* CLSI document EP15-A2 [ISBN 1-56238-574-7]. Clinical and Laboratory Standards Institute, 940 West Valley Road, Suite 1400, Wayne, Pennsylvania 19087-1898 USA, 2005.)

Proposed Guideline
December 1998

Approved Guideline
December 2001

Approved Guideline—Second Edition
June 2005

ISBN 1-56238-574-7
ISSN 0273-3099

Committee Membership

Area Committee on Evaluation Protocols

Luann Ochs, MS
Chairholder
Roche Diagnostics Corporation
Indianapolis, Indiana

Jan S. Krouwer, PhD
Vice-Chairholder
Krouwer Consulting
Sherborn, Massachusetts

George S. Cembrowski, MD, PhD
Provinicial Laboratory for Public Health
Edmonton, Alberta, Canada

David L. Duewer, PhD
National Institute of Standards and Technology
Gaithersburg, Maryland

Anders Kallner, MD, PhD
Karolinska Hospital
Stockholm, Sweden

Jacob (Jack) B. Levine, MBA
Bayer Corporation
Tarrytown, New York

Kristian Linnet, MD, PhD
Psychiatric University Hospital
Risskov, Denmark

Kristen L. Meier, PhD
FDA Center for Devices and Radiological Health
Rockville, Maryland

Daniel W. Tholen, MS
Dan Tholen Statistical Consulting
Traverse City, Michigan

Advisors

David A. Armbruster, PhD, DABCC, FACB
Abbott Laboratories
Abbott Park, Illinois

R. Neill Carey, PhD, FACB
Peninsula Regional Medical Center
Salisbury, Maryland

Carl C. Garber, PhD, FACB
Quest Diagnostics, Incorporated
Lyndhurst, New Jersey

Patricia E. Garrett, PhD
Boston Biomedica, Inc.
Portland, Maine

Martin H. Kroll, MD
Dallas VA Medical Center
Dallas, Texas

Donald M. Powers, PhD
Powers Consulting Services
Pittsford, New York

Max Robinowitz, MD
FDA Center for Devices and Radiological Health
Rockville, Maryland

Gian Alfredo Scassellati, PhD
Ente Nazional Italiano Di Unificatione
Turin, Italy

Jack Zakowski, PhD
Beckman Coulter, Inc.
Brea, California

Working Group on User Verification of Performance

R. Neill Carey, PhD, FACB
Chairholder
Peninsula Regional Medical Center
Salisbury, Maryland

F. Philip Anderson, PhD
Virginia Commonwealth University
Richmond, Virginia

Harvey George, PhD
MA Dept. of Public Health Laboratories
Jamaica Plain, Massachusetts

Alfred E. Hartmann, MD
Avera McKennan Hospital
Sioux Falls, South Dakota

Verlin K. Janzen, MD, FAAFP
Hutchinson Clinic, P.A.
Hutchinson, Kansas

Anders Kallner, MD, PhD
Karolinska Hospital
Stockholm, Sweden

Jacob (Jack) B. Levine, MBA
Bayer Corporation
Tarrytown, New York

Thomas J. Novicki, PhD, D(ABMM)
University of Washington
Seattle, Washington

Jennifer Schiffgens, MBA, MT(ASCP)
California Pacific Medical Center
San Francisco, California

Aparna Srinivasan, PhD
LifeScan, Inc.
Milpitas, California

Daniel W. Tholen, MS
Dan Tholen Statistical Consulting
Traverse City, Michigan

Staff

Clinical and Laboratory Standards Institute
Wayne, Pennsylvania

Lois M. Schmidt, DA
Staff Liaison

Donna M. Wilhelm
Editor

Melissa A. Lewis
Assistant Editor

The Area Committee on Evaluation Protocols would also like to recognize the original Clinical and Laboratory Standards Institute (CLSI) Subcommittee on User Demonstration of Performance who developed the first approved edition of this guideline.

R. Neill Carey, PhD, FACB, Chairholder
Harvey George, PhD
Alfred E. Hartmann, MD
Verlin K. Janzen, MD
Jacob B. Levine, MBA
W. Gregory Miller, PhD
Aparna Srinivasan, PhD

Contents

Contents (Continued)

Foreword

Before a laboratory can introduce a new method for reporting results of patient testing, several steps are required. First, the laboratory must specify the required performance for the method. Performance specifications may be defined by regulatory requirements and/or medical usefulness requirements. Second, the laboratory must select a method whose manufacturer's claims meet the required performance specifications. Finally, the laboratory must perform experiments to verify that the manufacturer's claimed imprecision and bias are achieved by the laboratory. If these steps are successful, the method is introduced into routine use for patient testing.

The focus of this guideline is verification of performance claims, for precision and trueness of a measurement procedure, that were previously validated by the manufacturer. This guideline is intended as a companion document to CLSI/NCCLS documents EP5—*Evaluation of Precision Performance of Quantitative Measurement Methods* and EP9—*Method Comparison and Bias Estimation Using Patient Samples*. EP5 and EP9 focus on the establishment and verification of performance claims. This document assumes that the manufacturer developed and validated performance claims using the protocols in EP5 and EP9. EP15 is intended to verify that a laboratory's performance is consistent with these claims.

The subcommittee had two principal goals during the development of EP15. One was to develop a testing protocol that is simple enough to be applicable in laboratories with a wide variety of sophistication and resources, from the point-of-care or physician's office laboratory to the large clinical laboratory. The second was to develop a protocol that is sufficiently rigorous to provide statistically valid conclusions for verification studies. To meet these two needs, the subcommittee developed a five-day testing protocol and simplified worksheets for all data gathering, statistical calculations, and tests of observed precision and trueness. A computer spreadsheet is provided to simplify and standardize the statistical calculations and tests of observed precision and trueness.

The first edition of EP15 used by clinical laboratories was judged by some to be difficult to understand for users who are not comfortable with statistics. EP15-A2 has removed the three-day protocol since it was determined that most methods did not qualify to use it. This protocol has fewer replicates than in EP15-A, and the spreadsheet should simplify calculations. Several terms have been changed to facilitate international harmonization (see below).

This document is primarily intended for use when an established method is initially set up in the laboratory. It may also be used to verify method performance after corrective action following a failed proficiency testing event.

A Note on Terminology

Clinical and Laboratory Standards Institute (CLSI) recognizes that harmonization of terms facilitates the global application of standards, and as a matter of organizational policy, is firmly committed to employing terms that are generally used internationally. This initiative includes a mechanism to resolve ISO/CEN/CLSI differences in nomenclature.

However, CLSI is also aware that medical conventions in the global metrological community have evolved differently in the United States, Europe, and elsewhere; that these differences are reflected in CLSI, ISO, and CEN documents; and that legally required use of terms, regional usage, and different consensus timelines are all obstacles to harmonization. Therefore, implementation of this policy must be an evolutionary and educational process that begins with new projects and revisions of existing documents.

The term *precision* is a measure of "closeness of agreement between independent test/measurement results obtained under stipulated conditions."[1] The terms in this document are consistent with uses

defined in the ISO 3534 and ISO 5725 series of standards. In these models, *repeatability* and *reproducibility* are considered to be the extreme measures of precision, with repeatability being the smallest measure (same operator, measurement procedure, equipment, time, and laboratory) and reproducibility being the largest (different operator, equipment, and laboratory). All other measures of precision are "intermediate measures" and must be explicitly described. Therefore, in this document, *within-run precision* has been replaced by *repeatability*. Reproducibility is not estimated since the EP15-A2 protocol does not require multiple laboratories. All other measures of precision from EP15-A have been retained, although the term *total precision* was eliminated because it is not clearly defined. In this document, *total precision* has been replaced by *within-laboratory precision*. Other harmonization changes include changing *specimen* to *sample* and *reportable range* to *measuring interval*.

Key Words

Bias, precision, repeatability, trueness, verification of performance

User Verification of Performance for Precision and Trueness; Approved Guideline—Second Edition

1 Scope

This guideline was developed for situations where the performance of the method was previously established and documented by experimental protocols with larger scope and duration. The experimental and statistical protocols of this guideline have relatively weak power to reject claims with statistical confidence, and should only be used to verify that the method is operating in accordance with the manufacturer's claims. This document is *not* intended to establish or validate the analytical performance of a method.

Since this document's scope is limited to verification of precision and trueness, other more rigorous CLSI/NCCLS protocols (e.g., EP6—*Evaluation of the Linearity of Quantitative Measurement Procedures: A Statistical Approach*; EP17—*Protocols for Determination of Limits of Detection and Limits of Quantitation*; and C28—*How to Define and Determine Reference Intervals in the Clinical Laboratory*) are employed to validate the method's performance against the user's needs. CLSI/NCCLS documents EP5—*Evaluation of Precision Performance of Quantitative Measurement Methods* and EP9—*Method Comparison and Bias Estimation Using Patient Samples* were developed to assist manufacturers in validating the performance of a diagnostic device for precision and trueness, respectively. CLSI/NCCLS document EP10—*Preliminary Evaluation of Quantitative Clinical Laboratory Methods* is intended for the rapid preliminary evaluation of precision, bias, sample carryover, drift, and nonlinearity. However, it is fairly complex since it is based on a multifactor design and is limited in the amount of data generated. EP10 should only be used as a preliminary evaluation of analytical performance.

One may also note that the EP15 protocol has an implicit assumption: namely, that if the estimated precision and trueness are acceptable, then the overall error (e.g., total analytical error) of the method is acceptable. This implied model can lead to an underestimation of the total analytical error[2] in cases where other effects are important. Besides conducting more extensive evaluations mentioned above, one could also consider performing the CLSI/NCCLS protocol EP21—*Estimation of Total Analytical Error for Clinical Laboratory Methods*. This protocol is a direct estimation of total analytical error and does not rely on a model.

2 Introduction

This guideline was written to assist the laboratory in verifying an established measurement procedure. It presumes that the procedure was checked by the manufacturer and is functioning properly. This guideline provides a minimum implementation protocol to verify that a particular example of a measurement procedure is operating in accordance with the manufacturer's claims. The laboratory must test the procedure against these targets for the protocols in this guideline to be applicable.

This guideline can also be used as a protocol to demonstrate acceptable performance when corrective actions are taken after failing proficiency testing (external quality assessment).

The specific characteristics addressed in this document are repeatability, within-laboratory precision, and trueness (as estimated by measures of bias) relative to an accepted standard. Upon successful completion of the protocols recommended in this guideline, the laboratory will have verified that the method is operating in accordance with the manufacturer's claims for precision and trueness.

This document leads the user through the process of determining the match between the laboratory's actual performance and the expected performance of the method. If the laboratory's performance is not consistent with the expected level of performance, remedial actions may be required.

Underlying this protocol is an assumption that the laboratory can operate the procedure properly.

3 Standard Precautions

Because it is often impossible to know what might be infectious, all patient and laboratory specimens are treated as infectious and handled according to "standard precautions." Standard precautions are guidelines that combine the major features of "universal precautions and body substance isolation" practices. Standard precautions cover the transmission of all infectious agents and thus are more comprehensive than universal precautions which are intended to apply only to transmission of blood-borne pathogens. Standard and universal precaution guidelines are available from the U.S. Centers for Disease Control and Prevention (*Guideline for Isolation Precautions in Hospitals*. Infection Control and Hospital Epidemiology. CDC. 1996;17(1):53-80 and *MMWR* 1988;37:377-388). For specific precautions for preventing the laboratory transmission of all infectious agents from laboratory instruments and materials and for recommendations for the management of exposure to all infectious disease, refer to the most current edition of CLSI/NCCLS document M29—*Protection of Laboratory Workers From Occupationally Acquired Infections*.

4 Definitions

analyte – component represented in the name of a measurable quantity (ISO 17511)[3]; **NOTE 1:** This includes any element, ion, compound, substance, factor, infectious agent, cell, organelle, activity (enzymatic, hormonal, or immunological), or property, activity, intensity, or other characteristics of which are to be determined; **NOTE 2:** In the type of quantity "mass of protein in 24-hour urine," "protein" is the analyte. In "amount of substance of glucose in plasma," "glucose" is the analyte. In both cases, the long phrase represents the **measurand** (ISO 17511)[3]; **NOTE 3:** In the type of quantity "catalytic concentration of lactate dehydrogenase isoenzyme 1 in plasma," "lactate dehydrogenase isoenzyme 1" is the analyte (ISO 18153).[4]

bias – the difference between the expectation of the test results and accepted reference value (ISO 3534-1).[5]

measurand – particular quantity subject to measurement (VIM93)[6]; **NOTE 1:** For example, the enzymatic activity of alkaline phosphatase at 37 °C; **NOTE 2:** The specification of a measurand may require statements about quantities such as time, temperature, and pressure (VIM93)[6]; **NOTE 3:** i.e., in the example above, the measurand includes not only the entity being measured (alkaline phosphatase), but the particular quality being measured (enzymatic activity), and the specific environmental condition under which it is being measured (37 °C).

measurement procedure – set of operations, described specifically, used in the performance of particular measurements according to a given method (VIM93)[6]; **NOTE 1:** This term pertains to specific procedures as marketed by specific manufacturers; **NOTE 2:** In other documents and in EP15-A, equivalent terms were **method**, **device**, and **assay**.

measuring interval – a set of values of measurands for which the error of a measuring instrument is intended to lie within specified limits; **NOTE 1:** "Error" is determined in relation to a conventional true value (VIM93)[6]; **NOTE 2:** the interval (or range) of values (in units appropriate for the analyte [measurand]) over which the acceptability criteria for the method have been met; that is, where errors due to nonlinearity, imprecision, or other sources are within defined limits; **NOTE 3:** Formerly, the term

reportable range was used in this document, and another commonly used term is **analytical measurement range**.

method of measurement – logical sequence of operations, described generically, used in the performance of measurements (VIM93)[6]; **NOTE 1:** A method of measurement, due to its generalized description, does not have numerically specified performance characteristics. A given method can be the basis of one or more measurement procedures, each with inherent numerical values for its performance characteristics (ISO 17511)[3]; **NOTE 2:** Formerly, the term **method** was used in EP15-A.

peer group – in proficiency testing, a group of presumably identical test systems.

precision (of measurement) – the closeness of agreement between independent test results obtained under stipulated conditions (ISO 3534-1);[5] **NOTE:** Precision is not typically represented as a numerical value but is expressed quantitatively in terms of **imprecision**—the standard deviation (SD) or the coefficient of variation (CV%) of the results in a set of replicate measurements.

repeatability (of results of measurements) – closeness of the agreement between the results of successive measurements of the same measurand carried out under the same conditions of measurement (VIM93)[6]; **NOTE:** Formerly, the term **within-run precision** was used in EP15-A.

repeatability conditions – conditions where independent test results are obtained with the same method on identical test material in the same laboratory by the same operator using the same equipment within a short interval of time (ISO 3534-1).[5]

run – an interval within which the trueness and precision of a testing system are expected to be stable, but cannot be greater than 24 hours. (U.S. CFR 493 February 28, 1992)[7]; **NOTE 1:** ISO 3534-1[5] defines "run" as follows: In a series of observations of a qualitative characteristic, the occurrence of an uninterrupted series of the same attribute is called a "run"; **NOTE 2:** Between analytical runs, events may occur that cause the measurement process to be susceptible to variations that are important.

total error – the sum of any set of defined errors that can affect the accuracy of an analytical result; **NOTE:** This document defines total error as the sum of bias and imprecision.

trueness (of measurement) – closeness of agreement between the average value obtained from a large series of test results and an accepted reference value (ISO 3534-1)[5]; **NOTE:** The measure of trueness is usually expressed in terms of bias (ISO 3534-1).[5]

within-laboratory precision – precision over a defined time and operators, within the same facility and using the same equipment. Calibration and reagents may vary; **NOTE:** Formerly, the term **total precision** was used in EP15-A.

5 Performance Standards

Prior to selecting a specific procedure for measuring an analyte and evaluating that procedure's performance, the laboratory must establish minimum performance specifications based on the clinical needs of the laboratory's clients. Lists of medically based performance standards are given in the references.[8-11] Some regulatory and accreditation programs[a] specify minimum standards for performance in proficiency testing. If regulatory performance standards apply, these define the maximum allowable measurement error the method can produce. These standards are expressed in terms of total allowable difference (total error) from an accepted reference value. Precision and trueness goals in terms of allowable standard deviation and bias must be derived from allowable total error. Discussions of the

[a] For example, in the U.S., CLIA and CAP.

relationship between allowable error and allowable standard deviation and bias are included in some of the publications listed in the references.[8-11] The user can also refer to the most current edition of CLSI/NCCLS document EP21—*Estimation of Total Analytical Error for Clinical Laboratory Methods*.

For the performance characteristics evaluated in this document, the following performance goal formats are recommended in order to conform to the evaluation result formats:

Precision. Precision goals should be stated as the maximum allowable SD and/or CV% at each analyte concentration to be tested.

Trueness. Goals for trueness should be stated as maximum allowable bias at each analyte concentration to be tested. Maximum allowable bias may be expressed as either an absolute concentration or as a percentage of the concentration.

Total Error. Total error goals should be stated as the maximum permissible difference between an individual sample's result and the target value for that sample. The target value may be determined by:

- the method's peer group in proficiency testing;
- an assigned reference method in proficiency testing;
- a comparative method in a comparison of a patient samples experiment; and
- the manufacturer of a reference material.

The user compares the manufacturer's claims to these performance goals. Ideally, the laboratory can select a method whose manufacturer's claims for precision and trueness are within the limits of the performance standards specified by the laboratory. If the manufacturer's claims are beyond the limits of the specified performance standards, this protocol is not appropriate. The user has the choice of validating the performance characteristics of the method using more extensive protocols such as those in CLSI/NCCLS documents EP5 and EP9, or selecting another candidate method, and comparing its claimed performance using the present protocol.

6 Overview of the Protocol

All of the experimental work in this protocol can be completed in five days.

6.1 Device Familiarization Period (see Section 7)

The device familiarization period is the time given to operators to become both familiar and comfortable with the details of the instrument's operation and the measurement procedure. Including a familiarization period into the timeline for an evaluation study is critical for meaningful evaluations of precision. If the operator has not had the opportunity for a familiarization period, including the opportunity to perform the measurement prior to beginning the precision protocol, the first data points generated by the operator may cause the laboratory to assume the test system has a higher level of imprecision and bias than is actually the case.

The familiarization period is also the time to verify that the QC materials the laboratory intends to use for the method perform as expected.

6.2 Precision Evaluation Experiment (see Section 8)

The precision evaluation experiment provides the user with guideline procedures for demonstrating precision performance. Usually, the manufacturer makes two types of precision claims—repeatability

(within-run precision) (σ_r) and within-laboratory precision (σ_l). This section provides statistical methods for identifying gross deviations from both types of claims.

6.3 Trueness Evaluation Experiment (see Section 9)

The trueness evaluation experiment provides the user with two different approaches. Either or both may be used.

(1) *Comparability.* Trueness may be assessed by way of a split-sample comparison experiment, by analyzing 20 patient samples distributed evenly over the entire measuring interval. Results from the two methods (the method under evaluation and a comparative method) are compared to determine if significant differences exist.

(2) *Recovery of expected values from certified reference materials.* Trueness may also be assessed by analyzing proficiency test materials and other assayed reference materials, and comparing the results for the method under evaluation to the expected reference value.

6.4 Measuring Interval and Reference Interval

User demonstration of measuring interval (range) is discussed in the most current edition of CLSI/NCCLS document EP6—*Evaluation of the Linearity of Quantitative Measurement Procedures: A Statistical Approach* (although the linear interval may not be the same as the measuring interval). User verification of reference intervals is included in the most current version of CLSI/NCCLS document C28—*How to Define and Determine Reference Intervals in the Clinical Laboratory*. These topics are not covered in this document.

7 Familiarization Period

After the system has been checked out by the manufacturer, staff must become familiar with the operation, maintenance procedures, methods of sample preparation, calibration, and monitoring functions. The length of time required for this process is variable, depending on the complexity of the device. Calibration should be verified during this period, if appropriate (see the most current edition of CLSI/NCCLS document EP6—*Evaluation of the Linearity of Quantitative Measurement Procedures: A Statistical Approach*). At the end of this time, the operator(s) should be confident in the operation of the device.

7.1 Operator Training

The operation, maintenance procedures, methods of sample preparation, and calibration and monitoring functions must be learned. Some manufacturers provide this training. The device should be set up and operated in the individual laboratory long enough to understand all of the procedures involved to avoid problems during the evaluation of its performance. Training should include the use of actual sample material, including pools, controls, leftover patient samples, or any other test materials appropriate for the device.

All possible contingencies (e.g., error flags, error correction, calibration) that may arise during routine operation should be carefully monitored. Data should not be collected during this period. Operator training is not complete until the user can confidently operate the device (see the most current edition of CLSI/NCCLS document GP21—*Training and Competence Assessment*).

7.2 Quality Control Procedures

Quality control procedures to be followed during the protocol are established during the familiarization period. It is important to verify that the device is operating in control, according to the manufacturer's specifications. To demonstrate this, use the control procedures recommended by the manufacturer. Due to the short duration of this protocol, the estimated standard deviations should not be used by themselves to establish quality control limits. For guidance on establishing ongoing quality control procedures, refer to the most current edition of CLSI/NCCLS document C24—*Statistical Quality Control for Quantitative Measurements: Principles and Definitions.*

7.3 Materials for Precision Experiments

Materials to be used as samples for precision experiments are tested during this period to verify that they perform as expected. Since precision may be directly related to concentration, analyte concentrations should be focused at or near medical decision points. For example, a glucose method's performance should be assessed at a concentration near 126 mg/dL (7.0 mmol/L), above which concentration, a fasting glucose result may indicate disease. For certain analytes, it may also be important to measure precision at the upper and lower limits of the measuring interval. Normally, it is sufficient to select materials that have analyte values near the concentrations the manufacturer used to establish the precision claims for the assay. (This information is in the package insert or instructions for use provided by the manufacturer of the test system under evaluation.) Acceptable materials for precision experiments include control samples (other than those used to assess whether the assay is in control), standards, previously analyzed patient samples, or suitable materials that have a known value. The materials used for precision experiments should mimic the matrix of the patient sample. For example, a measurement for a whole blood analyte should use a material for precision experiments that is as close as possible to human whole blood.

8 Verification of Precision Performance

Imprecision is a quantitative value indicating the extent of disagreement of a set of replicate measurements. Imprecision can be reported either as a standard deviation (SD) or a coefficient of variation (CV%), which expresses the standard deviation as a percentage of the mean value of the replicate measurements.

In either case, the mean value should be reported also. Increasing values of the standard deviation or coefficient of variation indicate increasing imprecision of the measurements.

In this document, precision is generally considered as either repeatability (within-run precision) or within-laboratory precision. Repeatability is a quantitative value indicating the disagreement among a set of replicate measurements when all measurements are made under identical conditions (or within a single run of a procedure).

Within-laboratory precision is a quantitative value indicating the disagreement among replicate measurements over a longer time when all known, major sources of measurement error within the laboratory (except for major maintenance, recalibration, or reagent lot changes) are accounted for. Within-laboratory precision reflects the accumulation of various error sources, including repeatability.

8.1 Experimental Design – Numbers of Days and Replicates

NOTE: The lower-case sigma (σ), combined with a subscript of "r" or "l" will be used to designate the manufacturer's claimed values of repeatability standard deviation and within-laboratory standard deviation, respectively (some manufacturers may call this "total precision"). The lower-case "s," combined with a subscript of "r" or "l" will be used to designate the user's estimated values of repeatability standard deviation and within-laboratory standard deviation, respectively.

8.2 Specific Procedures

(1) Analyze one run per day with three replicate samples at each of two concentrations daily for five days.

(2) If a run must be rejected because of quality control procedures or operating difficulties, discard the data, and conduct an additional run.

(3) Include the daily quality control samples normally used (see Section 7.2).

(4) Samples for the trueness experiment may be tested in the same runs (see Section 9).

(5) Calibrate as specified in the manufacturer's instructions for operators. If the manufacturer indicates in its claim that its precision data were generated over multiple calibration cycles, then the operator may choose to recalibrate during the experiment.

8.3 Recording the Data

Appendix A contains an example of a data recording sheet to summarize data. This type of summary is useful in the statistical analysis described below. An example of a completed worksheet is included in Appendix B. Alternatively, the user may wish to use the computer spreadsheet provided.

8.4 Calculation of Precision Estimates

After collecting the data and transcribing them onto an appropriate recording sheet, the calculations described in this section should be performed. A blank worksheet for the calculations is included in Appendix A. An example of a completed calculation worksheet is included in Appendix B. Alternatively, the user may wish to use the computer spreadsheet provided.

Separate calculations should be performed for each level of concentration.

8.4.1 Repeatability (Within-Run Precision)

Calculate the repeatability from the following formula (see the worksheet in Appendix A).

$$s_r = \sqrt{\frac{\sum_{d=1}^{D}\sum_{i=1}^{n}(x_{di} - \bar{x}_d)^2}{D(n-1)}}$$

where:

Σ indicates that the terms to the right of Σ are to be summed (see the worksheet in Appendix A),
D = total number of days (five),
n = total number of replicates per day (three),
x_{di} = result for replicates per day (three replicates), and
$\bar{x}_d$ = average of all results for day d.

It should be noted that if the experiment protocol is not followed exactly (same number of replicates for all runs on separate days), the repeatability estimate will be incorrect.

8.4.2 Within-Laboratory Precision

Within-laboratory precision is calculated using the following formulas (see worksheet in Appendix A). These calculations are based upon the variance components method discussed in CLSI/NCCLS document EP5—*Evaluation of Precision Performance of Quantitative Measurement Methods*.

Calculate the variance term, s_b^2, for the daily means from the formula (see the worksheet in Appendix A):

$$s_b^2 = \frac{\sum_{d=1}^{D} (\bar{x}_d - \bar{\bar{x}})^2}{D-1}$$

where:

$\bar{x}_d$ = average of all results for day d ($\bar{x}_1$ is average for day 1), and
$\bar{\bar{x}}$ = average of all results.

Calculate s_l from the formula (see the worksheet in Appendix A):

$$s_l = \sqrt{\frac{n-1}{n} \bullet s_r^2 + s_b^2}$$

where n = number of replicates per run (three).

8.5 Comparison of Estimated Repeatability to Manufacturer's Claims

Verify repeatability claims by comparing the repeatability estimate calculated in Section 8.4.1 to the manufacturer's claim. If the manufacturer's claim is in terms of coefficient of variation, convert the coefficient of variation into a standard deviation at the average concentration of all results for the material tested:

$$\sigma_r = CV\%_r \bullet \bar{\bar{x}}$$

where $CV\%_r$ is the manufacturer's claimed repeatability (within-run coefficient of variation).

If the estimated repeatability standard deviation is less than the manufacturer's claimed standard deviation, the user has demonstrated precision consistent with the claim. If the repeatability standard deviation is greater than the manufacturer's claim, note that a user's repeatability can be larger than the manufacturer's claim and not be statistically different from the claim. If the calculated repeatability standard deviation is larger than the manufacturer's claim, test whether it is statistically significantly larger (really different) as follows:

(1) Calculate the repeatability degrees of freedom, ν. For an experiment with D days duration and n replicates per run, ν is equal to D • (n-1). For the recommended protocol of five days duration and three replicates:

$$\nu = 10.$$

(2) Determine the $(1-\alpha/\ell)$ percentage point, C, of the χ^2 ("Chi-Square") distribution with ν degrees of freedom. Here, α is the false rejection rate (usually 5%), and ℓ is the number of levels tested. The percentage points for C corresponding to two, three, and four levels of testing are 97.5%, 98.33%, and 98.75%, respectively. Table 1 lists the values of C for these percentage points; other values can be

obtained from any standard statistics book or most commonly used computer spreadsheet programs. For the recommended protocols of five days duration, with two levels, C = 20.48.

(3) Calculate the verification value as

$$\frac{\sigma_r \bullet \sqrt{C}}{\sqrt{\nu}}$$

(4) If the estimated repeatability, s_r, is less than or equal to the verification value, data are consistent with the manufacturer's claim for repeatability, and the claim is verified. For better ability to detect departures from the manufacturer's claims, analyze two additional runs and recalculate all statistics, using the combined data. This will give a more powerful test of the manufacturer's claim (see Appendix C).

If the claim is not verified, then contact the manufacturer for assistance.

Table 1. Selected Percentage Points of the Chi-Square Distribution for Selected Numbers of Levels to Provide 5% False Rejection Rate

	Number of Levels		
Degree of Freedom	2	3	4
3	9.35	10.24	10.86
4	11.14	12.09	12.76
5	12.83	13.84	14.54
6	14.45	15.51	16.24
7	16.01	17.12	17.88
8	17.53	18.68	19.48
9	19.02	20.21	21.03
10	20.48	21.71	22.56
11	21.92	23.18	24.06
12	23.34	24.63	25.53
13	24.74	26.06	26.98
14	26.12	27.48	28.42
15	27.49	28.88	29.84
16	28.85	30.27	31.25
17	30.19	31.64	32.64
18	31.53	33.01	34.03
19	32.85	34.36	35.40
20	34.17	35.70	36.76
21	35.48	37.04	38.11
22	36.78	38.37	39.46
23	38.08	39.68	40.79
24	39.36	41.00	42.12
25	40.65	42.30	43.35

8.6 Comparison of Estimated Within-Laboratory Precision to Manufacturer's Claims

Verify within-laboratory claims by comparing the estimate calculated in Section 8.4.2 to the manufacturer's claim. If the manufacturer's claim for within-laboratory precision is in terms of coefficient of variation, convert the coefficient of variation into a standard deviation at the average concentration of all results for the material tested:

$$\sigma_l = CV\%_l \bullet \bar{\bar{x}}$$

where $CV\%_l$ is the manufacturer's claimed within-laboratory coefficient of variation (previous term, total coefficient of variation, CV_t).

If the estimated within-laboratory standard deviation is less than the manufacturer's claim, the user has demonstrated precision consistent with the claim. If the standard deviation is greater than the manufacturer's claim, note that a user's precision can be larger than the manufacturer's claim and not be statistically different from the claim. If the calculated repeatability standard deviation is larger than the manufacturer's claim, test whether it is statistically significantly larger (really different) as follows:

(1) Calculate the within-laboratory precision degrees of freedom, T. For an experiment with D days duration and n replicates per run (see the worksheet in Appendix A):

$$T = \frac{((n-1)s_r^2 + (ns_b^2))^2}{(\frac{n-1}{D})s_r^4 + (\frac{n^2(s_b^2)^2}{D-1})}$$

where s_b^2 is calculated in Section 8.4.2.

(2) Determine the $(1-\alpha/\ell)$ percentage point, C, of the χ^2 ("Chi-Square") distribution with ν degrees of freedom. Here, α is the false rejection rate (usually 5%), and ℓ is the number of levels tested. The percentage points for C corresponding to two, three, and four levels of testing are 97.5%, 98.33%, and 98.75%, respectively. Table 1 lists the values of C for these percentage points; other values can be obtained from any standard statistics book or most commonly used computer spreadsheet programs.

(3) Calculate the verification value as

$$\frac{\sigma_l}{T}\sqrt{C}$$

(4) If the within-laboratory precision estimate, s_l, is less than or equal to the verification value, the user has demonstrated precision consistent with the manufacturer's claim, and the claim can be considered as verified. If the user is not comfortable with this conclusion, then analyze two additional runs and recalculate all statistics using the combined data. This will give the user a more powerful test of the manufacturer's claims (see Appendix C).

If s_l exceeds the verification value, the user has not demonstrated precision consistent with the manufacturer's claim, and the user should contact the manufacturer for assistance.

9 Demonstration of Trueness

Trueness is conformance to a true value, accepted standard, or expected value. For a test result, bias is a measure of trueness; it is the difference between the test result and the accepted reference value for an

analyte. For a measurement procedure, bias is expressed as the difference between the average result obtained by a procedure under specified conditions and an accepted reference value, perhaps from an accepted comparative procedure or a certified reference material.

This protocol has provisions for demonstrating trueness by two procedures (see Appendixes D, E, and F; G and H):

(1) Comparison of patient sample results to another measurement procedure. When possible, the comparison of patient samples experiment should be performed. This technique avoids various artifacts, which can be present with commercial reference materials and some procedures. Comparison of patient samples is particularly important for initial evaluation of a procedure in a laboratory. It is strongly recommended when close agreement between procedures is expected. The comparison of patient samples can also form the basis of establishing the relationship between multiple procedures. This will ensure the laboratory's ability to provide equivalent results irrespective of which method is used (see the most current edition of CLSI/NCCLS document EP9—*Method Comparison and Bias Estimation Using Patient Samples*).

(2) Recovery of expected values for assayed reference materials. Once experience is gained with a procedure, the reference material strategy may be more convenient for verification following a calibration or proficiency test failure.

9.1 Comparison of Patient Sample Results to Those of Another Procedure

The choice of the comparative (reference) measurement procedure is critical for the interpretation of the results of this experiment.

If the laboratory intends to demonstrate trueness consistent with a manufacturer's claim, the comparative procedure must be the same as that used by the manufacturer in developing the claim. When the new procedure is a revision of a previous procedure from the same manufacturer, or application of the manufacturer's previous procedure to a new instrument, the laboratory's current procedure may be the manufacturer's comparative procedure. The manufacturer's trueness claim is applicable in this instance and should be used as the basis for demonstrating trueness in this experiment.

Often, the current procedure is different from the comparative procedure, or is a reference method or is performed in a reference laboratory, and the laboratory intends to demonstrate trueness of the new procedure relative to a procedure different from the one used by the manufacturer as the comparative procedure. In this case, the manufacturer's claim is not appropriate as the basis for demonstrating trueness, and should not be used as the basis for demonstrating trueness. The laboratory must specify a medically allowable bias between results obtained by the new procedure and the comparative procedure, and use this as the basis for demonstrating trueness. Guidance for specifying allowable bias is available.[8-12] The experimental protocol for comparing results of patient sample testing in this guideline has been designed as a verification protocol. To keep the experimental work simple, each patient is tested singly by the comparative and test procedures. Because this protocol can only detect relatively large bias between procedures, it should only be employed when the laboratory anticipates close agreement between the procedures (see Appendix F for more discussion). Otherwise, the laboratory should employ CLSI/NCCLS document EP9—*Method Comparison and Bias Estimation Using Patient Samples* because it includes more patient samples and testing in duplicate by both procedures. In all cases, it is important to verify that both procedures measure the same component of interest (have the same measurand).

Once a comparison procedure has been selected, the following steps should be followed. See the worksheet in Appendix D and the example of a completed worksheet in Appendix E. Alternatively, the user may wish to use the computer spreadsheet provided.

Obtain 20 patient samples in which the concentrations span but do not exceed the measuring interval of the measurement procedure. Do not use samples whose analyte concentrations exceed the measuring interval. The sample type must be compatible between the comparative and test procedures. Ideally, samples should be excluded that are known to contain substances identified as interferents in the manufacturer's instructions for either the test or comparative procedure (see the most current edition of CLSI/NCCLS document EP7—*Interference Testing in Clinical Chemistry*). This will be practical only in those cases where an EMR (electronic medical record) is in use and with the requisite software that allows for such an exclusion. If it is not practical to exclude these samples, see Section 9.1(2) which discusses discrepant results.

(1) Samples should be tested as close to the day of collection as possible to mimic conditions expected when the procedure is in regular use. If the comparative and test procedures are performed in geographically separate laboratories, it may be advisable to produce two sets of frozen aliquots. The aliquots can be thawed and analyzed at both laboratories around the same time, thus mitigating the issue of sample deterioration during transport. Follow the manufacturer's instructions for collection and handling of patients' samples. For laboratories in which abnormal samples are infrequently observed, it may be necessary to store them until there is a sufficient number for the experiment. If stored samples are used, they should be refrigerated, if consistent with stability of the measurand and manufacturer's instructions, to avoid possible artifacts introduced in the freeze-thaw cycle. Some laboratories may need to use frozen samples, particularly if the duration of the experiment and the manufacturer's product labeling require that frozen samples must be used, and they must be well mixed and examined for particulate matter after thawing and before use. If particulate matter is detected, they should be centrifuged, and the supernatant should be used for sampling. If stored samples must be used, testing by the comparative and test procedures should be performed within an hour or two if possible. The results of the measurements should be evenly spread over the measuring interval. If the full interval cannot be accommodated, the conclusions will only be applicable to the interval tested. A separate validation of the measuring interval may allow adequate conclusions regarding trueness and precision based on a restricted interval of sample results.

(2) Measure the samples on both the test and comparative procedures. These measurements should be completed within four hours of each other on the same day. Conclusions will be most reliable if the measurements are performed on five to seven samples per day for three to four days. This testing may be done concurrently with precision testing. Performing measurements on several days allows averaging of any between-day variability, which may exist for either measurement procedure. Examine the results after each event. If an isolated sample's results for the test and comparative procedures differ more than observed for other samples, then this is a signal that something may be wrong, and further investigation of the suitability of the assay may be necessary (i.e., beyond EP15). The following is recommended:

   (a) Measure that sample in duplicate on both measurement procedures. Some patient samples may show unexpectedly large differences due to differences in specificity between measurement procedures or for other reasons. Note that specificity-type errors will tend to be reproduced, but other random, systematic errors that are instrument- rather than patient-sample-related may not reproduce. Hence, it is unwise to conclude that a nonreproducible large difference is unimportant.

   (b) The discrepant data should be discarded for the EP15 analysis; otherwise, it will bias the results. To maintain the recommended sample size, one must add another sample to the analysis.

   (c) Use any information available in the original assay of the sample or in its reassay to find the root cause of the discrepancy. For example, a reproducible discrepancy suggests an interfering substance in the patient sample.

(3) Appropriate quality control procedures should be followed for each measurement procedure. Any unacceptable analytical performance should be corrected, and the samples from that run should be remeasured.

(4) Calculate the difference (or the individual sample bias) in reportable units and/or the percent difference (or percent individual sample percent bias) between each sample's results for the two procedures.

Individual sample bias in reportable units = b_i = (test procedure result$_i$ - comparison procedure result$_i$).

Individual sample bias in percent = $\%b_i$

$$\%b_i = 100 \bullet \left(\frac{\text{test procedure result}_i - \text{comparison procedure result}_i}{\text{comparison procedure result}_i} \right)$$

Construct a plot of bias and/or percent bias (vertical axis) vs. comparison procedure result (horizontal axis) for each sample. Examine the bias difference plot to determine if the difference between procedures is relatively constant over the concentration interval tested. If constant bias vs. concentration or constant percent bias difference vs. concentration is observed, then the mean bias in step (5) represents the average difference between the procedures. This value is compared to the manufacturer's claim to demonstrate the test measurement procedure's trueness.

If neither the bias nor the percent bias is constant over the concentration interval tested, the data should be partitioned into two segments and the average bias calculated separately for each segment. If the bias shows a progressively changing relationship to concentration, no average bias can be calculated. In this case, more data will be needed to validate the trueness of the test measurement procedure. Refer to the most current edition of CLSI/NCCLS document EP9—*Method Comparison and Bias Estimation Using Patient Samples* for additional information.

(5) Calculate the bias in reportable units and/or percent between the two procedures.

$$\bar{b} = \frac{\sum_{i=1}^{I} b_i}{n}$$

$$\overline{\%b} = \frac{\sum_{i=1}^{I} \%b_i}{n}$$

(6) Calculate the standard deviations of the bias and/or bias in percent.

$$s_{\bar{b}} = \sqrt{\frac{\sum_{i=1}^{I} (b_i - \bar{b})^2}{n-1}}$$

$$s_{\overline{\%b}} = \sqrt{\frac{\sum_{i=1}^{I} (\%b_i - \overline{\%b})^2}{n-1}}$$

9.1.1 Procedure for Demonstration of Trueness by Comparison of Patient Samples

If the bias or percent bias and the manufacturer's claimed bias or percent bias have the same sign, and the absolute value of the bias or percent bias is less than the absolute value of the manufacturer's claimed bias or percent bias, the laboratory has demonstrated bias consistent with the claim and no further statistical analysis is necessary. To further demonstrate trueness, the user may choose to use the procedures in Section 9.2.

If the absolute value of the bias or percent bias is greater than the absolute value of the manufacturer's claimed bias or percent bias, or if the bias has a different sign than the claim, then further statistical tests may be needed. Note that a laboratory's bias can be larger than the manufacturer's claim and not be statistically different from the claim. Test whether the bias or percent bias is statistically significantly larger (really different) than the manufacturer's claim using the procedure in Section 9.1.2.

This method can be used to test the bias only when the bias or percent bias is constant throughout the measuring interval, or to test the biases calculated by splitting the data as described in Section 9.1(4).

9.1.2 Acceptance Test for Demonstration of Trueness by Comparison of Patient Samples

Verification of the manufacturer's claim is performed using the following steps:

(1) Assume a false rejection rate, α. Typical values selected for this error rate are 1% and 5%.

(2) Determine the (100 - α) percent point, *t,* of the t-distribution with n-1 degrees of freedom. Here n represents the number of patient samples. For example, if α equals 1% and n equals 20, the (100 - α) point of the t-distribution with 19 degrees of freedom is 2.539. Other values of *t* can be obtained from any standard statistics book[13] or most commonly used computer spreadsheet programs for different values of α and n.

(3) Calculate the verification limits for bias in reportable units as

$$\beta - \frac{t \bullet s_{\bar{b}}}{\sqrt{n}} \quad \textbf{and} \quad \beta + \frac{t \bullet s_{\bar{b}}}{\sqrt{n}}$$

where β is the manufacturer's claimed value of bias.

If the estimated bias $\bar{b}$ is within the verification limits, the laboratory has demonstrated bias consistent with the manufacturer's claim.

If percent bias is used, calculate the verification limits for percent bias as

$$\beta - \frac{t \bullet s_{\overline{\%b}}}{\sqrt{n}} \quad \textbf{and} \quad \beta + \frac{t \bullet s_{\overline{\%b}}}{\sqrt{n}}$$

where β is the manufacturer's claimed value of percent bias.

If the estimated percent bias $\overline{\%b}$ is within the verification limits, the laboratory has demonstrated percent bias consistent with the manufacturer's claim.

(4) If the measured bias or percent bias is greater than the manufacturer's claim, but within the verification limits, the user may wish to make a more powerful test by measuring ten to 20 more patient samples and recalculating all statistics using the combined data.

(5) If the estimated bias is beyond the verification limits, the user has not demonstrated trueness consistent with the manufacturer's claim, and should contact the manufacturer for assistance.

9.2 Recovery of Expected Values From Reference Materials With Assigned Values

Reference materials with analyte target values are available from several sources. These materials are typically manufactured from human source materials but contain additives to achieve desired analyte concentrations and are processed to promote stability. Because of manufacturing requirements, processed materials have a solution matrix different from that of an authentic human sample. The difference in matrix may cause an altered analytical response, which is unique to a particular material-measurement procedure combination. Consequently, it is incorrect to use an analyte target value assigned by a definitive or reference procedure unless the reference material has been specifically evaluated and found suitable for use with the procedure being verified. Reference materials which have target values assigned specifically for the procedure of interest can be used to demonstrate trueness, provided that the target values are associated with known uncertainties. In this case, the trueness demonstration is limited to confirming that the test method performs similarly to that same procedure as used in other laboratories. Another consequence of material-procedure matrix interactions is the potential for different response with a lot of reagent not represented in the group of laboratories used for assignment of the target value. If a reagent lot difference occurs, the procedure-specific target value may not be appropriate for the new lot of reagents. Consult the manufacturer of the procedure for advice regarding appropriate reference materials for validation of trueness.

9.2.1 Sources of Reference Materials

Some sources of value-assigned materials, which can be used for verification of trueness, are listed below:

(1) Fresh frozen human serum or other unadulterated human materials. Certified Reference Materials (CRMs) for some analytes are available from the U.S. National Institute of Standards and Technology and other internationally recognized providers. A partial list of these materials is available from the Joint Committee for Traceability in Laboratory Medicine at: http://www.bipm.org/en/committees/jc/jctlm/jctlm-db/.

(2) Reference materials derived from proficiency testing programs. These materials are value assigned by a large number of laboratories and frequently represent numerous lots of reagents and system calibrators. Consequently, their target values represent average performance for the procedure.

(3) Materials provided by the method manufacturer for verification of trueness or quality control. These materials have been specifically designed for the measurement procedure being tested, but are generally not suitable for use with another manufacturer's measurement procedure.

(4) Materials used in interlaboratory quality control programs. These are measured by a relatively large number of laboratories, and their peer group mean values can be used as assigned values. Caution should be exercised that an adequate number of laboratories is included in the peer group for a reliable mean value. Ten participants in a procedure group is generally considered the minimum for a reliable mean value. These programs may include a small number of reagent lots for a procedure, which can affect the reliability of the target values for a new lot of reagents in an individual laboratory.

(5) Materials provided by third-party vendors, which have been value-assigned for a number of different measurement procedures. These are similar to proficiency testing or regional quality control materials, but generally have far fewer laboratories contributing to the peer group mean value. Consequently, the target value has a larger uncertainty. In addition, relatively few different lots of reagents may have been included, which can further affect the reliability of the assigned target values.

(6) Materials for which analyte concentrations can be fixed to stated concentrations, for example, partial gas pressures in blood can be fixed at stated values by tonometry.

9.2.2 Uncertainty of the Assigned Value

The manufacturer of the reference materials will often provide the uncertainty of the assigned values, or a 95% confidence interval for the values. For these materials, the user must determine the standard error of the assigned value in order to determine the verification limits (see Section 9.1.2). The standard error can be determined in a variety of ways, depending on the way the value was determined or the information provided by the manufacturer; for this document, call this value s_a:

(1) If the manufacturer provides the standard error or reports the "standard uncertainty" or "combined standard uncertainty" (u) of the assigned value, then this value is s_a (= u).

(2) If the manufacturer reports the "95% confidence interval" (CI) for the assigned value, then s_a = CI/2.

(3) If the manufacturer reports the "expanded uncertainty" of the assigned value (U), then it must also report the "coverage" (e.g., 95% or 99%) or the coverage factor. If coverage is 95%, s_a = U/2; if coverage is 99%, s_a = U/3. If the coverage factor is reported (often called "k"), then s_a = U/k.

(4) If the reference material assigned value comes from proficiency test consensus results, then the standard deviation of those results should also be reported (s), along with the number of results in the peer group (n). In this case, $s_a = (s/\sqrt{n})$.

(5) If the reference material assigned value comes from interlaboratory quality control programs, then s will depend on the information given in the program. Ideally, the standard deviation of results from different laboratories should be reported, along with the number of laboratories in the peer group (n). However, the program may report only the overall standard deviation of results (s) and the number of laboratories (n). In both of these cases, $s_a = (s/\sqrt{n})$, using the appropriate s and n as the number of laboratories. In the latter case, the s_a will be a slight overestimate of the uncertainty of the assigned value, but that is preferred to the underestimate that would occur if the total number of results was used as n.

9.2.3 Procedure for Demonstration of Trueness With Reference Materials

(1) Select the best available materials suitable for the measurement procedure. A minimum of two analyte concentrations should be tested, although more may be preferable to adequately evaluate the full measuring interval. The concentrations selected should represent the low and high ends of the interval for the measurement procedure. It may also be useful to test intermediate values corresponding to important medical decision levels. Use caution that the concentrations selected represent concentrations that can be measured with good precision by the test procedure.

(2) Prepare samples according to the manufacturer's instructions. Take precautions that the materials are thoroughly mixed prior to use.

(3) Measure each material in three to five different runs, each sample measured in duplicate.

(4) Calculate the mean ($\bar{x}$) and standard deviation (s_x) of the test results at each concentration.

9.2.4 Acceptance Test for Demonstration of Trueness With Reference Materials

Verification of the manufacturer's claim for trueness is performed using the following steps:

(1) Assume a false rejection rate, α. Typical values selected for this error rate are 1% and 5%.

(2) Assume the manufacturer has made no claim for bias relative to the assigned value (β=0).

(3) Determine the (100 - α) percent point, *t,* of the t-distribution with 2n-1 degrees of freedom. Here, n represents the number of samples tested and 2 represents the number of replicates (use 3 or 4 if that number of replicates is actually tested). For example, if α equals 1% and n equals 5, the (100 - α) point of the t-distribution with 9 degrees of freedom is 2.821. Other values of *t* can be obtained from any standard statistics book[13] or most commonly used computer spreadsheet programs for different values of α and n.

(4) Calculate the verification interval for bias in reportable units as

$$\bar{x} \pm t_{1-\alpha,\,2n-1} \bullet \sqrt{s_x^2 + s_a^2}$$

If percent bias is used, calculate the verification limits using percent bias and percent standard deviations.

(5) If the verification interval includes the assigned value, then the manufacturer's claim for trueness is verified.

(6) If the measured bias or percent bias is much different than the assigned value, but still within the verification interval, the user may wish to make a more powerful test by analyzing two to five more samples (in different runs) and recalculating all statistics using the combined data.

(7) If the assigned value is not included in the verification interval, the user has not demonstrated trueness consistent with the manufacturer's claim, and should consider one of the following options:

(a) Determine whether the bias and total error are acceptable for the laboratory's needs. Discussions of the relationship between allowable error and allowable standard deviation and bias are included in some of the publications listed in the references.[8-11] The user could also refer to the most current edition of CLSI/NCCLS document EP21—*Estimation of Total Analytical Error for Clinical Laboratory Methods*.

(b) Contact the manufacturer for assistance.

NOTE: With assigned values determined from proficiency testing, the degrees of freedom for the t statistic should consider the number of results used to determine the assigned value and s_a. However, this requires a level of statistical complexity beyond the scope of this simplified procedure.

References

[1] ISO. *Statistics – Vocabulary and Symbols – Part 2: Statistical Quality Control.* ISO 3534-2. Geneva: International Organization for Standardization; 1993.

[2] Krouwer JS. Setting performance goals and evaluating total analytical error for diagnostic assays. *Clin Chem*. 2002;48:919-927.

[3] ISO. In vitro *diagnostic medical devices – Measurement of quantities in biological samples – Metrological traceability of values assigned to calibrators and control materials*. ISO 17511. Geneva: International Organization for Standardization; 2003.

[4] ISO. In vitro *diagnostic medical devices – Measurement of quantities in biological samples – Metrological traceability of values assigned to catalytic concentration of enzymes in calibrators and control materials*. ISO 18153. Geneva: International Organization for Standardization; 2003.

[5] ISO. *Statistics – Vocabulary and Symbols – Part 1: Probability and General Statistical Terms*. ISO 3534-1. Geneva: International Organization for Standardization; 1993.

[6] ISO. *International Vocabulary of Basic and General Terms in Metrology*. Geneva: International Organization for Standardization; 1993.

[7] 42 CFR Part 493. *Medicare, Medicaid, and CLIA Programs; Laboratory Requirements Relating to Quality Systems and Certain Personnel Qualifications*; 1992.

[8] Garber CC, Carey RN. Evaluation of methods. In: Kaplan LA, Pesce AJ, eds. *Clinical Chemistry: Theory, Analysis, and Correlation*. 4th ed. St. Louis: Mosby; 2003:402-426.

[9] Koch DD, Peters T Jr. Selection and evaluation of methods. In: Burtis CA, Ashwood ER, eds. *Tietz Textbook of Clinical Chemistry*. 3rd ed. Philadelphia: W.B. Saunders Co; 1999:508-526.

[10] Petersen PH, Fraser CG, Kallner A, Kenny D. Strategies to set global analytical quality specifications in laboratory medicine. *Scand J Clin Lab Invest*. 1999;59:475-585.

[11] Fraser CG. *Biological Variation: From Principles to Practice.* Washington, DC: AACC Press; 2001.

[12] Westgard JO, Burnett RW. Precision requirements for cost-effective operation of analytical processes. *Clin Chem.* 1990;36:1629-1632.

[13] ASQC. *Glossary & Tables for Statistical Quality Control*. Milwaukee, WI: ASQC Quality Press; 1983:1160.

Appendix A. Sample Data Recording Sheet – Precision Experiment

Use a separate sheet for each concentration.

Device _____ Analyte _____ Concentration _____

Reagent Source/Lot _____ Calibrator Source/Lot _____

	Run 1	Run 2	Run 3	Run 4	Run 5
Date/Operator					
Replicate 1 (x_1)					
Replicate 2 (x_2)					
Replicate 3 (x_3)					
$\sum_{i=1}^{3} x_i = x_1 + x_2 + x_3$					
$\bar{x}_d = \frac{\sum_{i=1}^{3} x_i}{3}$					
$x_1 - \bar{x}_d$					
$(x_1 - \bar{x}_d)^2$					
$x_2 - \bar{x}_d$					
$(x_2 - \bar{x}_d)^2$					
$x_3 - \bar{x}_d$					
$(x_3 - \bar{x}_d)^2$					
$\sum_{1}^{3} (x_i - x_d)^2 = (x_1 - x_d)^2 + (x_2 - x_d)^2 + (x_3 - x_d)^2$					
$sd_{run}^2 = \frac{\sum_{i=1}^{3} (x_i - \bar{x}_d)^2}{2}$					
$\bar{x}_d - \bar{\bar{x}}$ (see below)					
$(\bar{x}_d - \bar{\bar{x}})^2$					

Appendix A. (Continued)

1. Calculation of Grand Mean $= \bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \bar{x}_4 + \bar{x}_5}{5} =$ _____

2. Calculation of s_r :

$$sd^2_{run,\ average} = \frac{sd^2_{run\ 1} + sd^2_{run\ 2} + sd^2_{run\ 3} + sd^2_{run\ 4} + sd^2_{run\ 5}}{5} = \_\_\_\_\_$$

$$s_r = \sqrt{sd^2_{run,\ average}} = \_\_\_\_\_$$

3. Calculation of s_l:

s_r = _____ , from above

$$s_b^2 = \frac{\sum_{d=1}^{5} \left(\bar{x}_d - \bar{\bar{x}}\right)^2}{4} = \frac{(\bar{x}_1 - \bar{\bar{x}})^2 + (\bar{x}_2 - \bar{\bar{x}})^2 + (\bar{x}_3 - \bar{\bar{x}})^2 + (\bar{x}_4 - \bar{\bar{x}})^2 + (\bar{x}_5 - \bar{\bar{x}})^2}{4} = \_\_\_\_\_$$

$$s_l = \sqrt{\frac{n-1}{n} \bullet s_r^2 + s_b^2}$$, where n = number of replicates in each run.

$$s_l = \sqrt{\frac{2}{3} \bullet s_r^2 + s_b^2} = \_\_\_\_\_$$

4. Verification of Repeatability Claim:

Compare calculated s_r to claimed σ_r:

If calculated s_r < claimed σ_r, repeatability has been demonstrated to be consistent with the manufacturer's claim. See paragraph 5, "Verification of Within-Laboratory Claim."

If calculated s_r > claimed σ_r, note that a user's repeatability can be larger than the manufacturer's claim and not be statistically different from the claim.

Calculate the verification value

$\nu = 10$, $C = 20.48$

Verification value $= \frac{\sigma_r \bullet \sqrt{C}}{\sqrt{\nu}} = \sigma_r \bullet 1.431 =$ _____

Appendix A. (Continued)

Compare calculated s_r to verification value:

If calculated $s_r \leq$ verification value, repeatability has been demonstrated to be consistent with the manufacturer's claim.

If calculated $s_r >$ verification value, repeatability has not been demonstrated to be consistent with the manufacturer's claim; contact the manufacturer for help.

5. Verification of Within-Laboratory Claim:

Compare calculated s_l to claimed σ_l:

If calculated $s_l <$ claimed σ_l, within-laboratory precision has been demonstrated to be consistent with the manufacturer's claim.

If calculated $s_l >$ claimed σ_l, note that a user's within-laboratory precision can be larger than the manufacturer's claim and not be statistically different from the claim.

Calculate degrees of freedom, T:

$$T = \frac{((n-1)s_r^2 + (ns_b^2))^2}{(\frac{n-1}{D})s_r^4 + (\frac{n^2(s_b^2)^2}{D-1})},$$

where D = number of days, and
n = number of replicates.

$$T = \frac{(2s_r^2 + 3(s_b^2))^2}{0.4s_r^4 + \frac{9B^2}{4}} = \_\_\_\_\_$$

C = _____ (obtain from Table 1 for T degrees of freedom)

$$\text{Verification value} = \frac{\sigma_l \bullet \sqrt{C}}{\sqrt{T}} = \_\_\_\_\_$$

Compare calculated s_l to verification value:

If calculated $s_l \leq$ verification value, within-laboratory precision has been demonstrated to be consistent with the manufacturer's claim.

If calculated $s_l >$ verification value, within-laboratory precision has not been demonstrated to be consistent with the manufacturer's claim; contact the manufacturer for help. Additional information that may help with troubleshooting include the date of the last calibration, the actual results from the controls from each run, and the expected ranges of the controls and other information that is provided with the controls.

Appendix B. Example of a Completed Sample Data Recording Sheet – Precision Experiment

Device XYZ Analyte Glucose Concentration 140 mg/dL

Reagent Source/Lot MK243 Calibrator Source/Lot RNC59YR

Assume the manufacturer's claim for σ_r is 1.0 mg/dL and for σ_l is 2.0 mg/dL.

	Run 1	Run 2	Run 3	Run 4	Run 5
Date/Operator	2/20 TF	2/21 JL	2/22 GG	2/23 KW	2/24 SR
Replicate 1 (x_1)	140	138	143	143	142
Replicate 2 (x_2)	140	139	144	143	143
Replicate 3 (x_3)	140	138	144	142	141
$\sum_{i=1}^{3} x_i = x_1 + x_2 + x_3$	420	415	431	428	426
$\bar{x}_d = \frac{\sum_{i=1}^{3} x_i}{3}$	140.00	138.33	143.67	142.67	142.00
$x_1 - \bar{x}_d$	0	-0.33	-0.67	0.33	0.00
$(x_1 - \bar{x}_d)^2$	0	0.1089	0.4489	0.1089	0.0000
$x_2 - \bar{x}_d$	0	0.67	0.33	0.33	1.00
$(x_2 - \bar{x}_d)^2$	0	0.4489	0.1089	0.1089	1.0000
$x_3 - \bar{x}_d$	0	-0.33	0.33	-0.67	-1.00
$(x_3 - \bar{x}_d)^2$	0	0.1089	0.1089	0.4489	1.0000
$\sum_{1}^{3} (x_i - x_d)^2 = (x_1 - x_d)^2 + (x_2 - x_d)^2 + (x_3 - x_d)^2$	0	0.6667	0.6667	0.6667	2.0000
$sd^2_{run} = \frac{\sum_{i=1}^{3} (x_i - \bar{x}_d)^2}{2}$	0	0.3334	0.3334	0.3334	1.0000
$\bar{x}_d - \bar{\bar{x}}$ (see below)	1.33	3.00	2.34	1.34	0.67
$(\bar{x}_d - \bar{\bar{x}})^2$	1.7689	9.000	5.4756	1.7956	0.4489

1. Calculation of Grand Mean = $\bar{\bar{x}}$

$$\bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \bar{x}_4 + \bar{x}_5}{5} = \frac{140.00 + 138.33 + 143.67 + 142.67 + 142.00}{5} = \underline{141.33}$$

2. Calculation of s_r:

$$sd^2_{run,\ average} = \frac{sd^2_{run\ 1} + sd^2_{run\ 2} + sd^2_{run\ 3} + sd^2_{run\ 4} + sd^2_{run\ 5}}{5} =$$

Appendix B. (Continued)

$$sd^2_{run,\ average} = \frac{0.0000 + 0.3334 + 0.3334 + 0.3334 + 1.0000}{5} = \underline{0.40004}$$

$$s_r = \sqrt{sd^2_{run,\ average}} = \underline{0.632}$$

3. Calculation of s_l (within-laboratory standard deviation)

$s_r = \underline{0.632}$, from above

$$s_b^2 = \frac{\sum_{d=1}^{5} (\bar{x}_d - \bar{\bar{x}})^2}{4} = \frac{(\bar{x}_1 - \bar{\bar{x}})^2 + (\bar{x}_2 - \bar{\bar{x}})^2 + (\bar{x}_3 - \bar{\bar{x}})^2 + (\bar{x}_4 - \bar{\bar{x}})^2 + (\bar{x}_5 - \bar{\bar{x}})^2}{4} = \underline{4.62225}$$

$s_l = \sqrt{\frac{n-1}{n} \bullet s_r^2 + s_b^2}$, where n = number of replicates in each run.

$$s_l = \sqrt{\frac{2}{3} \bullet s_r^2 + s_b^2} = \sqrt{\frac{2}{3} \bullet 0.4004 + 4.62225} = \underline{2.21 \text{ mg/dL}}$$

4. Verification of Repeatability Claim:

Compare calculated s_r to claimed σ_r:

If calculated s_r < claimed σ_r, repeatability has been demonstrated to be consistent with the manufacturer's claim. See paragraph 10, "Verification of Within-Laboratory Precision Claim."

If calculated s_r > claimed σ_r, note that a user's repeatability can be larger than the manufacturer's claim and not be statistically different from the claim.

Calculate the verification value

$\nu = 10$, $C = 20.48$

$$\text{Verification value} = \frac{\sigma_r \bullet \sqrt{C}}{\sqrt{\nu}} = \sigma_r \bullet 1.431 = 1.0 \bullet 1.431 \text{ mg/dL} = \underline{1.431 \text{ mg/dL}}$$

Compare calculated s_r to verification value:

If calculated $s_r \leq$ verification value, repeatability has been demonstrated to be consistent with the manufacturer's claim.

Appendix B. (Continued)

If calculated s_r > verification value, repeatability has not been demonstrated to be consistent with the manufacturer's claim; contact the manufacturer for help.

s_r = 0.63 mg/dL, which is less than the manufacturer's claim, and less than the verification value; repeatability has been verified to be consistent with the manufacturer's claimed repeatability.

5. Verification of Within-Laboratory Precision Claim:

Compare calculated s_l to claimed σ_l:

If calculated s_l < claimed σ_l, within-laboratory precision has been demonstrated to be consistent with the manufacturer's claim.

If calculated s_l > claimed σ_l, note that a user's within-laboratory precision can be larger than the manufacturer's claim and not be statistically different from the claim.

Calculate degrees of freedom, T:

$$T = \frac{((n-1)s_r^2 + (ns_b^2))^2}{(\frac{n-1}{D})s_r^4 + (\frac{n^2(s_b^2)^2}{D-1})},$$

where D = number of days, and
n = number of replicates.

$$T = \frac{(2s_r^2 + 3s_b^2)^2}{0.4s_r^4 + \frac{9(s_b^2)^2}{4}} = \frac{((2 \bullet 0.632^2) + (3 \bullet 4.62225))^2}{0.4 \bullet 0.632^4 + \frac{9 \bullet 4.2225^2}{4}} = \frac{(0.8008 + 13.86675)^2}{(0.010281 + 48.07169)} = \frac{215.137}{48.082} = \underline{4.47}$$

C = 11.14 (assumed 2 levels) (obtain from Table 1 for T degrees of freedom)

$$\text{Verification value} = \frac{\sigma_l \bullet \sqrt{C}}{\sqrt{T}} = \frac{2.0 \bullet \sqrt{11.14}}{\sqrt{4.47}} = \underline{3.16\,\text{mg/dL}}$$

Compare calculated s_l to verification value:

If calculated $s_l \leq$ verification value, within-laboratory precision has been demonstrated to be consistent with the manufacturer's claim.

If calculated s_l > verification value, within-laboratory precision has not been demonstrated to be consistent with the manufacturer's claim; contact the manufacturer for help.

s_l = 2.21 mg/dL, which exceeds the manufacturer's claim, but is less than the verification value; within-laboratory precision has been verified to be consistent with the manufacturer's claimed within-laboratory precision.

Appendix C. Additional Statistical Explanations and Considerations – Precision Experiment

The appropriate experimental design to employ for demonstrating that a user's calculated precision is consistent with a manufacturer's precision claims depends upon the following four factors:

- the statistical false rejection (Type I or "alpha" error) rate;
- the statistical false acceptance (Type II or "beta" error) rate;
- the number of analyte concentrations for which imprecision claims are being verified; and
- the size of the claimed repeatability, σ_r, relative to the claimed within-laboratory precision, σ_l.

Changes in each of these four factors may affect the number of days duration of the experiment. In the experimental design recommended in this document, one run per experimental day is required. Consequently, the total number of runs is equal to the number of days duration of the experiment. It is assumed that claims are being verified for all analyte concentrations and that the same design is used for all levels.

The statistical false rejection testing rate is the chance that the user will falsely conclude that the manufacturer's claims are incorrect. Typical values selected for this error rate are 1% and 5%, and a false rejection rate of 5%. As the false rejection rate decreases, the required amount of collected data increases. This additional amount of information could result in a longer experiment (that is, more days).

The statistical false acceptance testing rate is the chance that the user will falsely conclude that the manufacturer's claims are correct. Typical values selected for this error rate are 1%, 5%, and 10%. As the false acceptance rate decreases, the required amount of collected data increases. False acceptance rates are reported below for both of the recommended experimental designs. If multiple runs per day are performed with either of these experimental designs, then the statistical false acceptance rate will be smaller than the reported rates.

The number of analyte concentrations being tested indirectly impacts the experimental design through the false rejection rate. Without the proper mathematical corrections, as the number of analyte concentrations increases, so does the overall false rejection rate. For example, if a 5% false rejection rate is utilized for each level, when four analyte concentrations are considered, the chance of falsely rejecting at least one claim exceeds 18%. Mathematically, an overall false rejection rate of a specified size, such as the 5% rate associated with both recommended experimental designs, is achieved by lowering the false rejection rate for each level. Unfortunately, lowering the single-level false rejection rate increases the required amount of data to detect a difference between the observed precision and the manufacturer's claimed precision with the same degree of power.

The size of σ_r relative to σ_l affects both the required number of replicates per run and the duration of the experiment. If σ_r is relatively small when compared to σ_l, then the number of replicates per run is less important than the total number of runs. In contrast, if σ_r is relatively large when compared to σ_l, then the number of replicates per run is relatively more important than the experimental duration. If σ_r and σ_l are approximately equal, then the number of replicates per run and the number of runs are about of equal importance.

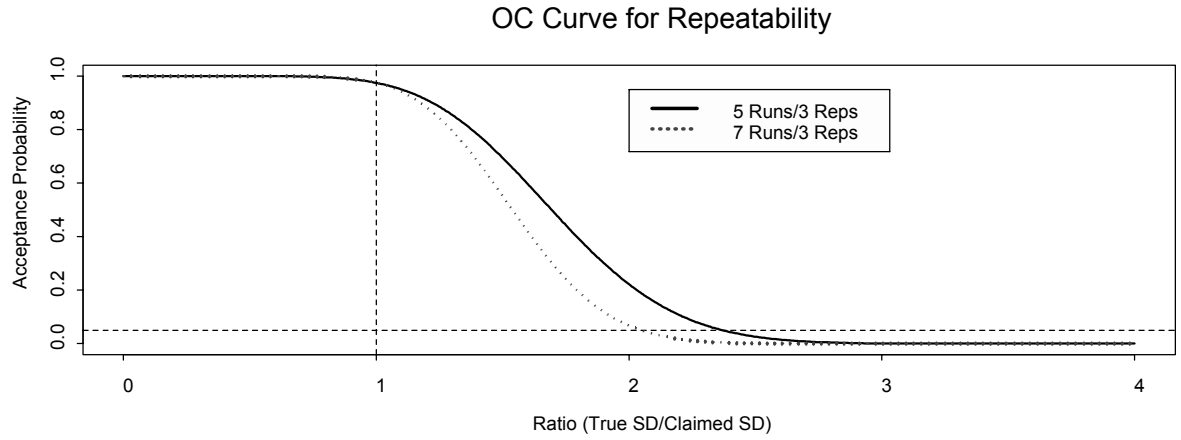
The size of σ_r relative to σ_l affects the degrees of freedom (obtained using Satterthwaite approximation) for the test design. For the purpose of determining the power of the recommended test, the degrees of freedom is obtained assuming that σ_r is relatively small when compared to σ_l. This results in a conservative estimate (fewer degrees of freedom) of the total degrees of freedom for within-laboratory estimate.

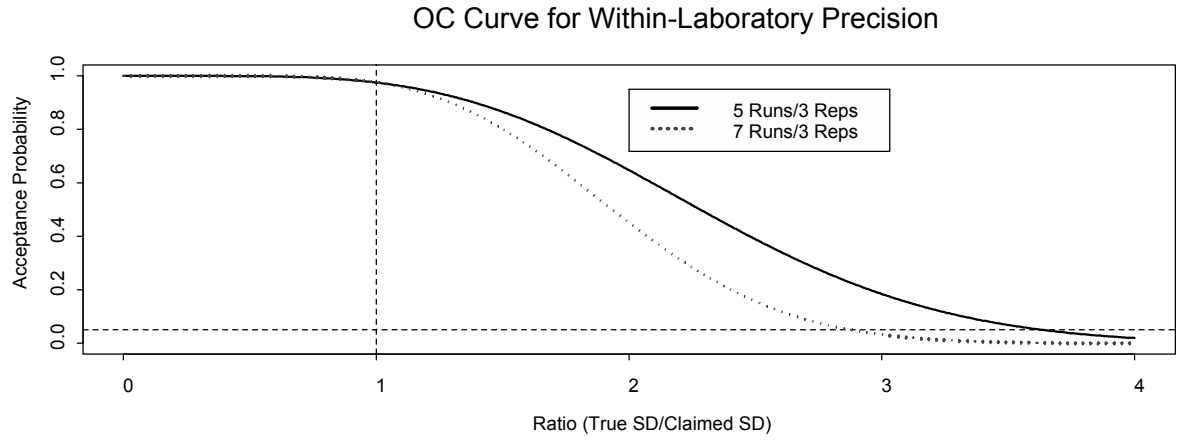
Appendix C. (Continued)

The operating characteristic (OC) curve, also called the power curve or risk curve, can help determine the optimal test design to balance risk and cost. This curve is a plot of the probability of accepting the hypothesis (the hypothesis being tested is that the repeatability or within-laboratory precision = manufacturer's claimed value), versus the true state of nature. The following plots depict the OC curves for the following two test designs for testing both repeatability and within-laboratory precision.

- Five days, one run per day, and three replicates per run
- Seven days, one run per day, and three replicates per run

In both of these designs, a false rejection rate of 5% is assumed if precision is tested at two concentrations. The effects of increasing the number of concentrations on the performance characteristics of the experiment are discussed below.





Appendix C. (Continued)

For the prescribed design of five days, one run per day, and three replicates per run tested at two concentrations, the above OC curve is interpreted as follows:

- For repeatability, the probability of accepting the hypothesis is greater than 95% when the ratio of the true SD to the claimed value is less than 1 and the probability of accepting the hypothesis is less than 5% when the ratio is greater than 2.5.

- For within-laboratory precision, the probability of accepting the hypothesis is greater than 95% when the ratio of the true SD to the manufacturer's claimed value is less than 1 and the probability of accepting the hypothesis is less than 5% when the ratio is greater than 3.5.

When the number of days is increased from five to seven, the probability of accepting the hypothesis for repeatability is less than 5% when the ratio of the true SD to the manufacturer's claimed value is greater than 2. If the number of analyte levels increases, the amount of data to achieve the same probability of acceptance increases.

The OC curves show the risk levels for the test designs considered. As can be seen, the proposed test design can only detect relatively large deviations from claims. This further emphasizes that this protocol should be applied only in situations in which the performance of the method is already very well characterized.

If lower risk levels are desired, the user may increase the number of replicates within each run, if σ_r is relatively large compared to σ_l, or increase the number of days tested if σ_l is relatively large compared to σ_r. Similarly, if the number of analyte levels tested is greater than two, the user may wish to increase either the number of replicates or the number of days tested.

Appendix D. Sample Data Recording Sheet – Comparison of Patient Samples Experiment (see equations for statistical terms on next page)

Test Method Result	Comp. Method Result	b_i	$b_i - \bar{b}$	$\left(b_i - \bar{b}\right)^2$	$\%b_i$	$\%b_i - \overline{\%b}$	$\left(\%b_i - \overline{\%b}\right)^2$
Sums							

Appendix D. (Continued)

Equations for calculating terms in the data recording sheet:

b_i = (test procedure result$_i$ - comparison procedure result$_i$)

$$\bar{b} = \frac{\sum_{i=1}^{I} b_i}{n}$$

Individual sample bias in percent = $\%b_i$

$$\%b_i = 100 \bullet \left(\frac{\text{test procedure result}_i - \text{comparison procedure result}_i}{\text{comparison procedure result}_i} \right)$$

1. Calculation of mean bias in reportable units and percent between the two procedures:

$$\bar{b} = \frac{\sum_{i}^{I} b_i}{n} = \_\_\_\_\_$$

$$\overline{\%b} = \frac{\sum_{i}^{I} \%b_i}{n} = \_\_\_\_\_$$

2. Calculation of standard deviations of the bias and percent bias:

$$s_{\bar{b}} = \sqrt{\frac{\sum_{i}^{I} (b_i - \bar{b})^2}{n-1}} = \_\_\_\_\_$$

$$s_{\overline{\%b}} = \sqrt{\frac{\sum_{i}^{I} (\%b_i - \overline{\%b})^2}{n-1}} = \_\_\_\_\_$$

3. Calculation of verification value for bias in reportable units and percent bias:

- Do the calculated bias ($\bar{b}$) or percent bias ($\overline{\%b}$) and the manufacturer's claimed bias or percent bias have the same sign? Yes/No

- Is the absolute value of the bias (or percent bias) less than the absolute value of the manufacturer's claimed bias? Yes/No

Appendix D. (Continued)

If the answer to both questions is "Yes," the laboratory has demonstrated bias consistent with the claim.

If the answer to either question is "No," note that a user's bias can be larger than the manufacturer's claim and not be statistically different from the claim. The observed bias or percent bias must be compared with the appropriate verification limits.

Verification limits (bias) = $\beta - \frac{t \bullet s_{\bar{b}}}{\sqrt{n}} =$ _____ and $\beta + \frac{t \bullet s_{\bar{b}}}{\sqrt{n}} =$ _____ .

Verification limits (percent bias) = $\beta - \frac{t \bullet s_{\overline{\%b}}}{\sqrt{n}} =$ _____ and $\beta + \frac{t \bullet s_{\overline{\%b}}}{\sqrt{n}} =$ _____ .

Compare calculated bias or percent bias to verification limits. If calculated bias or percent bias is within the verification limits, bias or percent bias has been demonstrated to be consistent with the manufacturer's claim. If bias or percent bias exceeds appropriate verification value, bias has not been demonstrated to be consistent with the manufacturer's claim; contact the manufacturer.

Appendix E. Example of a Completed Sample Data Recording Sheet – Comparison of Patient Samples Experiment

Assume the manufacturer's claim for bias is 2.0 mg/dL for a glucose method at 126 mg/dL [7.00 mmol/L].

Test Procedure Result	Comp. Procedure Result	b_i	$b_i - \bar{b}$	$(b_i - \bar{b})^2$	$\%b_i$	$\%b_i - \overline{\%b}$	$(\%b_i - \overline{\%b})^2$
76	77	-1	-3.5	12.25	-1.30	-3.66	13.39
127	121	6	3.5	12.25	4.96	2.60	6.75
256	262	-6	-8.5	72.25	-2.29	-4.65	21.63
303	294	9	6.5	42.25	3.06	0.70	0.49
29	25	4	1.5	2.25	16.00	13.64	186.02
345	348	-3	-5.5	30.25	-0.86	-3.22	10.39
42	41	1	-1.5	2.25	2.44	0.08	0.01
154	154	0	-2.5	6.25	0.00	-2.36	5.57
398	388	10	7.5	56.25	2.58	0.22	0.05
93	92	1	-1.5	2.25	1.09	-1.27	1.62
240	239	1	-1.5	2.25	0.42	-1.94	3.77
72	69	3	0.5	0.25	4.35	1.99	3.95
312	308	4	1.5	2.25	1.30	-1.06	1.13
99	101	-2	-4.5	20.25	-1.98	-4.34	18.85
375	375	0	-2.5	6.25	0.00	-2.36	5.57
168	162	6	3.5	12.25	3.70	1.34	1.80
59	54	5	2.5	6.25	9.26	6.90	47.59
183	185	-2	-4.5	20.25	-1.08	-3.44	11.85
213	204	9	6.5	42.25	4.41	2.05	4.21
436	431	5	2.5	6.25	1.16	-1.20	1.44
Sums		50		357	47.22		346.08

Appendix E. (Continued)

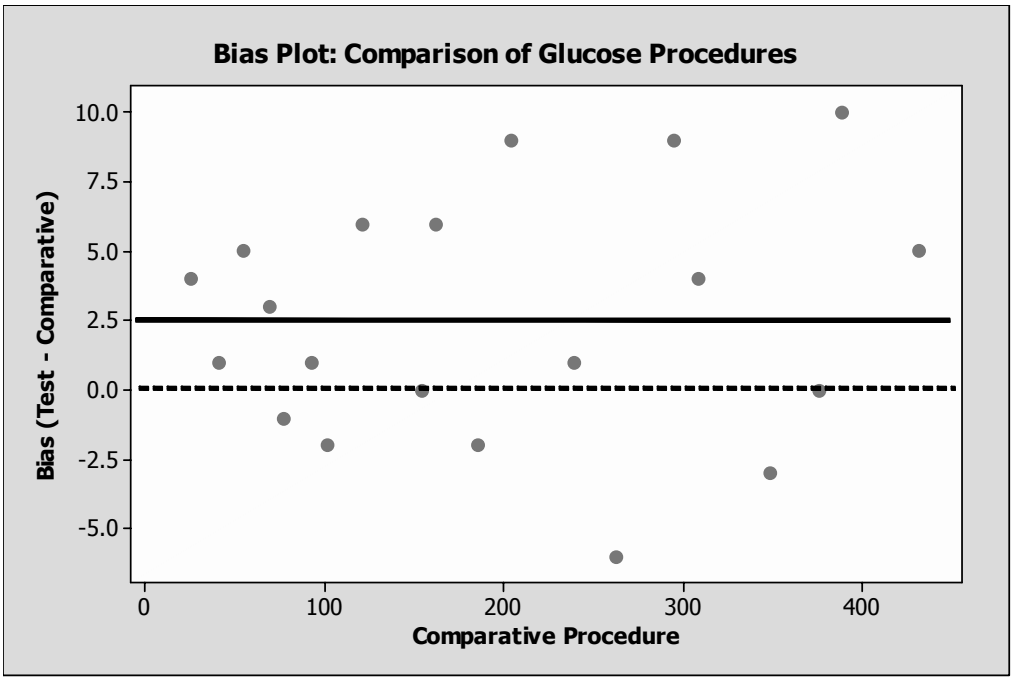


Figure E1. Bias Plot of Sample Data From Comparison of Patient Samples Experiment. The bias is indicated by the solid line at 2.50 mg/dL. The ideal bias (zero) is indicated by the dashed line at zero. Note that the bias is approximately constant over the concentration interval of the experiment.

1. Calculation of mean bias in reportable units and percent between the two procedures:

$$\bar{b} = \frac{\sum_{i}^{I} b_i}{n} = \frac{50}{20} = \underline{2.50\,\text{mg/dL}}$$

$$\overline{\%b} = \frac{\sum_{i}^{I} \%b_i}{n} = \frac{47.22}{20} = \underline{2.36\%}$$

2. Calculation of standard deviations of the bias and percent bias:

$$s_{\bar{b}} = \sqrt{\frac{\sum_{i}^{I} (b_i - \bar{b})^2}{n-1}} = \sqrt{\frac{357.00}{19}} = \underline{4.33\ \text{mg/dL}}$$

$$s_{\overline{\%b}} = \sqrt{\frac{\sum_{i}^{I} (\%b_i - \overline{\%b})^2}{n-1}} = \sqrt{\frac{346.08}{19}} = \underline{4.27\%}$$

Appendix E. (Continued)

3. Calculation of verification value for bias in reportable units and percent bias:

   - Do the calculated bias ($\bar{b}$) or percent bias ($\overline{\%b}$) and the manufacturer's claimed bias or percent bias have the same sign? Yes

   - Is the absolute value of the bias (or percent bias) less than the absolute value of the manufacturer's claimed bias? No

   If the answer to both questions is "Yes," the laboratory has demonstrated bias consistent with the claim.

   If the answer to either question is "No," note that a user's bias can be larger than the manufacturer's claim and not be statistically different from the claim. The observed bias or percent bias must be compared with the appropriate verification limits.

   Verification limits (bias) =

$$\beta - \frac{t \cdot s_{\bar{b}}}{\sqrt{n}} = 2.00\,\text{mg/dL} - \frac{2.539 \cdot 4.33\,\text{mg/dL}}{\sqrt{20}} = \underline{\quad 0.46\,\text{mg/dL} \quad} \text{ and}$$

$$\beta + \frac{t \cdot s_{\bar{b}}}{\sqrt{n}} = 2.00\,\text{mg/dL} + \frac{2.539 \cdot 4.33\,\text{mg/dL}}{\sqrt{20}} = \underline{4.46\,\text{mg/dL}}$$

   The calculated bias (2.50 mg/dL) is within the verification limits; bias has been shown to be consistent with the manufacturer's claim.

Appendix F. Power Calculation for the Procedure for Verification of Trueness by Comparison of Patient Samples

In this procedure, we are essentially testing the following hypothesis:

$H_{0:}$ $b \leq b_0$ (where b_0 is the manufacturer's claimed bias)
versus
$H_{1:}$ $b > b_0$

In the procedure, the two levels of alpha considered are 5% and 1% (or 95% and 99% confidence levels, respectively).

Background:

The power curve or the OC (operating characteristic) curve can help determine the optimal number of samples that balances cost and risk.

Risks:

Passing the test when the bias is different from the claimed value is called "producers risk." In statistical terms, this is Type II error, which has probability β.

NOTE: Do not confuse this with the β used in the guideline for the manufacturer's claimed bias value. Here, b_0 represents the manufacturer's claimed bias value.

Failing the test when the bias is the same as the claimed value is called "consumers risk." In statistical terms, this is Type I error, which has probability α.

Cost vs. Risk:

Small number of samples = lower cost, higher risk
Large number of samples = higher cost, lower risk

A number of factors influence power:

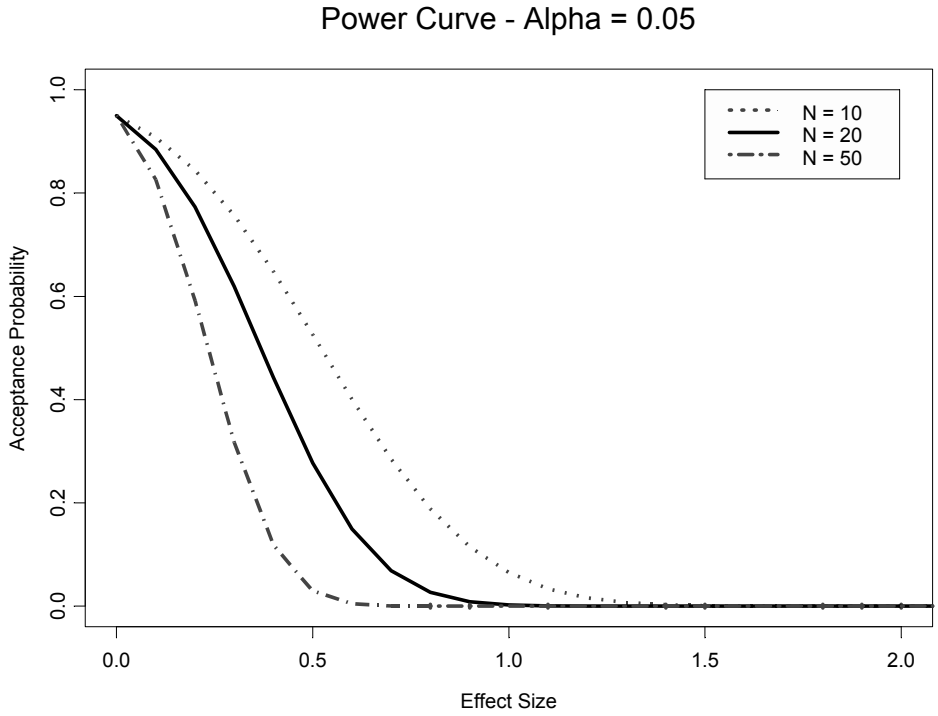
1. α, the probability of a Type I error (level of significance). As the probability of a Type I error (α) increases, the probability of a Type II error (β) decreases. Hence, as α increases, power = 1 - β also increases.

2. σ, the variability in the population. As σ increases, power decreases.

3. The size of the population difference (delta) (i.e., the difference between the claimed response [true] and the estimated value, expressed as a fraction [or multiple] of the repeatability standard deviation). As the effect decreases, power decreases. In the following plots, the power is computed as a function of delta/σ, called the effect size. This represents the number of standard deviations the claimed bias is away from the estimated bias.

4. Number of samples. As the number of samples increases, power increases.

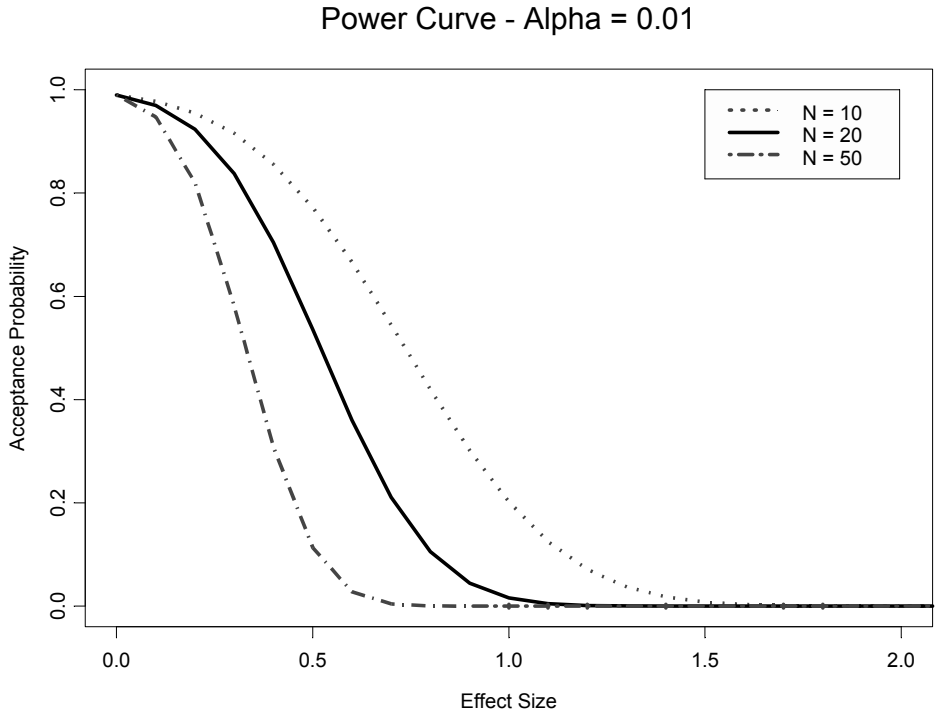
Appendix F. (Continued)



The power curves for different numbers of samples are presented below for alpha = 0.05 and 0.01, respectively. As can be seen from Graph 1 (alpha = 0.05), as long as the number of samples is at least 10, the probability of accepting the null hypothesis is small when the estimated average bias is 1 SD away from the claimed value. For the 20 samples recommended in this guideline (solid black line), the null hypothesis is accepted with a probability:

1. of less than 0.3 (30%) when the estimated average bias is ½ SD away from the claimed value; and
2. of less than 0.05 (5%) when the estimated average bias is ¾ SD away from the claimed value.

Appendix F. (Continued)



When alpha = 0.01, at least 20 samples are required to pick up a difference of ¾ SD between the claimed value and an estimated average bias with high probability (low probability of acceptance = high probability of detection).

Appendix G. Sample Data Recording Sheet – Demonstration of Trueness With Reference Materials

Use a separate sheet for each concentration. More runs and replicates are optional.

Device _____ Analyte _____ Assigned Concentration _____

Reagent Source/Lot _____ Calibrator Source/Lot _____

	Test Result, x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
Run 1, Date/Operator	-	-	-
Replicate 1 (x_1)			
Replicate 2 (x_2)			
Run 2, Date/Operator	-	-	-
Replicate 1 (x_3)			
Replicate 2 (x_4)			
Run 3, Date/Operator	-	-	-
Replicate 1 (x_5)			
Replicate 2 (x_6)			
Run 4, Date/Operator	-	-	-
Replicate 1 (x_3)			
Replicate 2 (x_4)			
Run 5, Date/Operator	-	-	-
Replicate 1 (x_5)			
Replicate 2 (x_6)			
$\sum_{i=1}^{n} x_i = (x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10})$		-	-
$\bar{x} = \frac{\sum_{i}^{n} x_i}{n} = \frac{\sum_{1}^{6} x_i}{10}$		-	-
$\sum_{i=1}^{n}(x_i - \bar{x})^2 = (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + (x_4 - \bar{x})^2 + (x_5 - \bar{x})^2 + (x_6 - \bar{x})^2 + (x_7 - \bar{x})^2 + (x_8 - \bar{x})^2 + (x_9 - \bar{x})^2 + (x_{10} - \bar{x})^2$	-	-	

1. Calculation of mean:

$$\bar{x} = \frac{\sum_{i}^{n} x_i}{n} = \frac{(x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10})}{10} = \_\_\_\_\_$$

Appendix G. (Continued)

2. Calculation of standard deviation:

$$s_x^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{\sum_{1}^{10}(x_i - \bar{x})^2}{9} =$$

$$s_x^2 = \frac{(x_1-\bar{x})^2 + (x_2-\bar{x})^2 + (x_3-\bar{x})^2 + (x_4-\bar{x})^2 + (x_5-\bar{x})^2 + (x_6-\bar{x})^2 + (x_7-\bar{x})^2 + (x_8-\bar{x})^2 + (x_9-\bar{x})^2 + (x_{10}-\bar{x})^2}{9}$$

$$s_x = \sqrt{s_x^2} = \_\_\_\_\_$$

3. Calculate verification interval for bias in reportable units.

Verification interval $= \bar{x} \pm t_{1-\alpha,2n-1} \bullet \sqrt{s_x^2 + s_a^2}$, with s_a defined or calculated according to Section 9.2.2.

Determine the (100 - α) percent point, *t,* of the t-distribution with 2n-1 degrees of freedom. Here, n represents the number of samples tested and 2 represents the number of replicates (use 3 or 4 if that number of replicates is actually tested). For example, if α equals 1% and n equals 10, the (100 - α) point of the t-distribution with[13] 9 degrees of freedom is 2.821. Other values of *t* can be obtained from any standard statistics book or most commonly used computer spreadsheet programs for different values of α and n.

1. If the verification interval includes the assigned value, then the manufacturer's claim for trueness is verified.

2. If the measured bias or percent bias is much different than the assigned value, but still within the verification interval, the user may wish to make a more powerful test by analyzing two to five more samples (in different runs) and recalculating all statistics using the combined data.

3. If the assigned value is not included in the verification interval, the user has not demonstrated trueness consistent with the manufacturer's claim, and should consider one of the following options:

- Determine whether the bias and total error are acceptable for the laboratory's needs. Discussions of the relationship between allowable error and allowable standard deviation and bias are included in some of the publications listed in the references.[7-11]

- Contact the manufacturer for assistance. Additional information that may help with troubleshooting include the date of the last calibration, the actual results from the reference materials from each run, and the expected ranges of the reference materials and other information that is provided with the reference materials.

Appendix H. Example of a Completed Sample Data Recording Sheet – Demonstration of Trueness With Reference Materials

Use a separate sheet for each concentration. More runs and replicates are optional.

Device XYZ Analyte Glucose Assigned Concentration 40 mg/dL

Reagent Source/Lot MK243 Calibrator Source/Lot RNC59YR

	Test Result, x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
Run 1, Date/Operator 2/20 TF	-	-	-
Replicate 1 (x_1)	37	-0.7	0.49
Replicate 2 (x_2)	38	0.3	0.09
Run 2, Date/Operator 2/21 JL	-	-	-
Replicate 1 (x_3)	39	1.3	1.69
Replicate 2 (x_4)	37	-0.7	0.49
Run 3, Date/Operator 2/22 GG	-	-	-
Replicate 1 (x_5)	38	0.3	0.09
Replicate 2 (x_6)	36	-1.7	2.89
Run 4, Date/Operator 2/23 KW	-	-	-
Replicate 1 (x_3)	39	1.3	1.69
Replicate 2 (x_4)	38	0.3	0.09
Run 5, Date/Operator 2/24 SR	-	-	-
Replicate 1 (x_5)	38	0.3	0.09
Replicate 2 (x_6)	37	-0.7	0.49
$\sum_{i=1}^{n} x_i = (x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10})$	377	-	-
$\bar{x} = \frac{\sum_{i}^{n} x_i}{n} = \frac{\sum_{1}^{6} x_i}{10}$	37.7	-	-
$\sum_{i=1}^{n}(x_i - \bar{x})^2 = (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + (x_4 - \bar{x})^2 + (x_5 - \bar{x})^2 + (x_6 - \bar{x})^2 + (x_7 - \bar{x})^2 + (x_8 - \bar{x})^2 + (x_9 - \bar{x})^2 + (x_{10} - \bar{x})^2$	-	-	8.1

1. Calculation of mean:

$$\bar{x} = \frac{\sum_{i}^{n} x_i}{n} = \frac{(x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10})}{10} = \frac{377}{10} = \underline{37.7 \text{ mg/dL}}$$

Appendix H. (Continued)

2. Calculation of standard deviation:

$$s_x^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{\sum_{1}^{10}(x_i - \bar{x})^2}{9} =$$

$$s_x^2 = \frac{(x_1-\bar{x})^2 + (x_2-\bar{x})^2 + (x_3-\bar{x})^2 + (x_4-\bar{x})^2 + (x_5-\bar{x})^2 + (x_6-\bar{x})^2 + (x_7-\bar{x})^2 + (x_8-\bar{x})^2 + (x_9-\bar{x})^2 + (x_{10}-\bar{x})^2}{9}$$

$$s_x^2 = \frac{8.1}{9}$$

$$s_x = \sqrt{s_x^2} = \underline{0.949 \text{ mg/dL}}$$

3. Calculate verification interval for bias in reportable units.

   Verification interval = $\bar{x} \pm t_{1-\alpha,2n-1} \bullet \sqrt{s_x^2 + s_a^2}$, with s_a defined or calculated according to Section 9.2.2.

   Determine the (100 - α) percent point, *t,* of the t-distribution with 2n-1 degrees of freedom. Here, n represents the number of samples tested and 2 represents the number of replicates (use 3 or 4 if that number of replicates is actually tested). For example, if α equals 1% and n equals 5, the (100 - α) point of the t-distribution with 9 degrees of freedom is 2.821. Other values of *t* can be obtained from any standard statistics book[13] or most commonly used computer spreadsheet programs for different values of α and n.

   Degrees of freedom = 2n-1 (2 replicates, n samples); since there were 5 runs, n = 5, and <u>degrees of freedom = 9.</u>

   <u>t = 2.821</u> (9 degrees of freedom, α = 1%).

   Assume the material was a proficiency testing sample, with a peer group standard deviation of 1.73 mg/dL, and a peer group of 135 laboratories. Then,

$$s_B = \frac{s_{\text{peer group}}}{\sqrt{n}} = \frac{1.73 \text{ mg/dL}}{\sqrt{135}} = \frac{1.73 \text{ mg/dL}}{11.62} = \underline{0.149 \text{ mg/dL}}$$

$$\text{Verification interval} = \bar{x} \pm t_{1-\alpha,2n-1} \bullet \sqrt{s_x^2 + s_a^2}$$

$$\text{Verification interval} = 37.7 \pm 2.821\sqrt{0.949^2 + 0.149^2} = 37.7 \pm 2.821\sqrt{0.923}$$

$$\text{Verification interval} = 37.7 \pm 2.71 = \underline{34.99 - 40.41 \text{ mg/dL}}$$

4. If the verification interval includes the assigned value, then the manufacturer's claim for trueness is verified.

Appendix H. (Continued)

The verification interval (34.99 to 40.41 mg/dL) includes the assigned value (40 mg/dL).

5. If the measured bias or percent bias is much different than the assigned value, but still within the verification interval, the user may wish to make a more powerful test by analyzing two to five more samples (in different runs) and recalculating all statistics using the combined data.

6. If the assigned value is not included in the verification interval, the user has not demonstrated trueness consistent with the manufacturer's claim, and should consider one of the following options:

   a. Determine whether the bias and total error are acceptable for the laboratory's needs. Discussions of the relationship between allowable error and allowable standard deviation and bias are included in some of the publications listed in the references.[7-11]

   b. Contact the manufacturer for assistance. Additional information that may help with troubleshooting include the date of the last calibration, the actual results from the reference materials from each run, and the expected ranges of the reference materials and other information that is provided with the reference materials.

Appendix I. Flow Sheet for Experiments for User Demonstration of Performance for Trueness and Precision (Accuracy)

I. PURPOSE: This guideline is designed to verify that a procedure is performing within the manufacturer's previously validated performance claims for precision and trueness (accuracy).

II. DEFINITIONS

precision (of measurement) – the closeness of agreement between independent tests results obtained under stipulated conditions (ISO 3534-1)[5]; **NOTE:** Precision is not typically represented as a numerical value but is expressed quantitatively in terms of imprecision—the standard deviation (SD) or the coefficient of variation (CV%) of the results in a set of replicate measurements.

trueness – closeness of agreement between the average value obtained from a large series of test results and an accepted reference value; **NOTE 1:** The measure of trueness is usually expressed in terms of bias; **NOTE 2:** Trueness has previously been termed accuracy.

bias – the difference between the expectation of the test results and accepted reference value (ISO 3534-1)[5]; **NOTE:** Numerical expression of trueness estimated as the difference between the average result obtained by a method under specified conditions and an accepted reference value (e.g., a "conventional true value," from an accepted comparative method or a certified reference material).

run – interval within which the trueness and precision of a testing system are expected to be stable, but cannot be greater than 24 hours.

III. MATERIALS REQUIRED

A. Precision Experiment – two samples with sufficient quantity for 15 analyses each.

1. Acceptable materials include control samples (other than those used to assess whether the assay is in control), standards, previously run patient samples, or suitable materials that have a known value.

2. **NOTE:** The materials used for precision experiments should mimic the matrix of the patient sample (i.e., use a material that is as close as possible to human whole blood, serum, or plasma). Select materials that have analyte values near the concentrations the manufacturer used to establish the precision claims for the assay. (This information is in the package insert or instructions for use provided by the manufacturer of the test system under evaluation.) Ideally, commercial quality controls or reference materials should be available which fall within +/-10% of the concentration levels listed in the precision claims in the package insert. If these are not readily available, the user can create mini-pools within these intervals by appropriate mixtures of serum pools, controls, or reference materials. Since the pools will only approximate the mean value of the precision data, comparison of the CV% at these levels is a better choice than the SD.

B. Trueness Experiment: This protocol has two options for demonstrating trueness. Sample requirements depend on the option selected.

Appendix I. (Continued)

1. Comparison of patient sample results to another method. If the laboratory intends to demonstrate trueness consistent with a manufacturer's claim, the comparative procedure must be the same as that used by the manufacturer in developing the claim. If this comparison method is not available, then option two is recommended.

    a) Samples – 20 or more samples in which the concentration of the selected analyte concentrations span but do not exceed the analytical interval of the assay. The sample type must be compatible between the comparative and test procedures. Follow the manufacturer's instructions for collection and handling of patients' samples. If stored samples are used, they should be stored consistent with the manufacturer's instructions. If stored samples must be used, testing by the comparative and test procedures should be performed within an hour or two if possible.

2. Recovery of expected values for assayed reference materials. Two or more samples in which the concentration of the selected analyte concentrations spans but does not exceed the analytical measuring interval of the assay. Possible materials include:

    a) Fresh frozen human serum or other nonadulterated human materials. Such materials are available from National Institute of Standards and Technology (NIST) with NRSCL reference and definitive values assigned for some analytes.

    b) Reference materials derived from proficiency testing programs. These materials are value-assigned by a large number of laboratories and frequently represent numerous lots of reagents and system calibrators.

    c) Materials provided by the method manufacturer for verification of trueness or quality control. These materials have been specifically designed for the measurement procedure being tested, but are generally not suitable for use on another manufacturer's method.

    d) Materials used in interlaboratory quality control programs. These materials are measured by a relatively large number of laboratories, and their peer group mean values can be used to assess agreement.

IV. PROCEDURE

A. Notes

1. All of the experimental work in this protocol can be completed in five days.

2. Operators must be familiar with the device. Before running this protocol, the device should be set up and operating in the individual laboratory long enough to ensure that training and competency are not a factor in the performance of the evaluation. Data should not be collected during the training period.

3. Quality control procedures are established during the familiarization period. It is important to verify that the device is operating in accordance with the manufacturer's specifications. To demonstrate this, use the control procedures recommended by the manufacturer during this protocol.

Appendix I. (Continued)

B. Demonstration of Precision

1. Protocol

a) Analyze two sample concentrations three times each day for five days.

b) Include QC samples in each day's run. If QC is not in control, reject that run and discard the data. Correct the problem and conduct an additional run.

c) Record the data on the data recording sheet in Appendix A.

1) If using the Appendix A data recording sheet, perform calculations as indicated and follow instructions contained therein.

d) If using the computer spreadsheet, follow instructions provided with the spreadsheet.

C. Demonstration of Trueness (can be run concurrently with precision experiment)

Select experiment: comparison of patient sample results, or recovery of expected values from reference materials.

1. Comparison of Patient Sample Results to Another Method

a) Ideally, the comparative method should be the same method used by the manufacturer as the reference method in the claim. If not, see Section 9.1.

b) Obtain 20 patient samples whose analyte concentrations are evenly distributed throughout the analytical measuring interval (reportable range) of the test method.

c) Measure the samples on the test and comparative methods within four hours of each other if possible. Spread the testing out over the five days of the precision experiment.

d) Record the data on the data recording sheet in Appendix D.

1) If using the Appendix D data recording sheet, perform calculations as indicated and follow instructions contained therein.

e) If using the computer spreadsheet, follow instructions provided with the spreadsheet.

2. Recovery of Expected Values From Reference Materials

a) Select the best available materials to be tested (see Section 9.2).

b) Prepare materials according to the manufacturer's instructions.

c) Measure each sample with two or more replicates in three to five runs.

d) Record the data on the data recording sheet in Appendix G.

e) If using the Appendix G worksheet, perform calculations as indicated and follow instructions contained therein.

f) If using the computer spreadsheet, follow instructions provided with the spreadsheet.

CLSI consensus procedures include an appeals process that is described in detail in Section 8 of the Administrative Procedures. For further information, contact CLSI or visit our website at www.clsi.org.

Summary of Consensus Comments and Committee Responses

EP15-A: *User Demonstration of Performance for Precision and Accuracy; Approved Guideline*

General

1. I want to express my concerns about terminology. I do not think we can accept a philosophy like that expressed on page 4. Namely, is it acceptable in a CLSI guideline to call "B" what VIM calls "A" (e.g., accuracy instead of trueness; analytical method instead of measurement procedure; and total precision instead of reproducibility)? CLSI guidelines still continue to ignore the correct VIM terminology or to use terms in open contrast with it. This is not harmonization, and certainly is far from globalization.

- **CLSI recognizes that harmonization of terms facilitates the global application of standards, and as a matter of organizational policy, is firmly committed to employing terms that are generally used internationally. The terminology in EP15-A2 has been harmonized with ISO 3534-1: 1993- *Statistics - Vocabulary and Symbols – Part 1: Probability and General Statistical Terms;* VIM:1993 – *International Vocabulary, Basic and General Terms in Metrology;* and ISO 3534-2: 1993 – *Statistics – Vocabulary and Symbols – Part 2: Statistical Quality Control.***

2. I attached the Excel-sheet that we use for the evaluation of comparisons (See abstract F-7 on page A174 of the June 2002 edition of Clinical Chemistry). It is largely self-explaining. The graphs on the sheet 'differences' will be displayed only if you enter a character into cells r1 or r2 and the allowable error into column B. The workbook I sent is for five components only; we have had reasons to work with up to 25 comparisons and have a version for that available.

- **The working group will consider this approach in the development of computer spreadsheets that perform all calculations in EP15-A2.**

Appendix A, Sample Data Recording Sheets – Precision Experiment

3. Calculation of Grand Mean and B (page 18, equations 1 and 3, respectively). Does $\bar{x}_1$ found in those equations $= \bar{x}_d$, run 1 and $\bar{x}_2 = \bar{x}_d$ run 2, and so forth? If not, please explain their meaning.

- **The Sample Data Recording Sheet of EP15-A2 has been revised to clarify the calculation of the grand mean.**

Appendix C, Additional Statistical Explanations and Considerations – Precision Experiment

4. In Section 5.1, the second paragraph states: "Formulas to determine the appropriate experimental designs to verify imprecision claims for more than two analyte levels or other error rates are given in Appendix C," but there are no such formulas in Appendix C.

- **All formulae are now included in Sections 8.4 to 8.6.**

Summary of Consensus/Delegate Comments and Committee Responses

EP15-A2: *User Verification of Performance for Precision and Trueness; Approved Guideline—Second Edition*

General

1. The calculation worksheets tend to be cumbersome; Excel spreadsheets could be made available on the CLSI website for use. The spreadsheet could do the calculations.

- **Companion software for EP15-A2 is under development.**

Appendix A, Sample Data Recording Sheet – Precision Experiment

2. On page 19, in the 12th line of the table, xi should read x3!

- **The reviewer is correct. The table entry in Appendix A has been corrected.**

Appendix A, Sample Data Recording Sheet – Precision Experiment; and Appendix G, Sample Data Recording Sheet – Demonstration of Trueness With Reference Materials

3. Additional information mentioned in Sections 7.1 and 7.2 that could be included on these sheets to assist in troubleshooting are:

   1) The date of last calibration
   2) The actual results of the controls from each run
   3) The expected ranges of the controls (or info attached).

- **The reviewer's suggested comments were added to Appendices G and H.**

NOTES

The Quality System Approach

Clinical and Laboratory Standards Institute (CLSI) subscribes to a quality management system approach in the development of standards and guidelines, which facilitates project management; defines a document structure via a template; and provides a process to identify needed documents. The approach is based on the model presented in the most current edition of CLSI/NCCLS document HS1—*A Quality Management System Model for Health Care.* The quality management system approach applies a core set of "quality system essentials" (QSEs), basic to any organization, to all operations in any healthcare service's path of workflow (i.e., operational aspects that define how a particular product or service is provided). The QSEs provide the framework for delivery of any type of product or service, serving as a manager's guide. The quality system essentials (QSEs) are:

Documents & Records	Equipment	Information Management	Process Improvement
Organization	Purchasing & Inventory	Occurrence Management	Service & Satisfaction
Personnel	Process Control	Assessment	Facilities & Safety

EP15-A2 addresses the quality system essentials (QSEs) indicated by an "X." For a description of the other documents listed in the grid, please refer to the Related CLSI/NCCLS Publications section on the following page.

Documents & Records	Organization	Personnel	Equipment	Purchasing & Inventory	Process Control	Information Management	Occurrence Management	Assessment	Process Improvement	Service & Satisfaction	Facilities & Safety
					X C24 C28 EP5 EP6 EP7 EP9 EP10 GP10						

Adapted from CLSI/NCCLS document HS1—*A Quality Management System Model for Health Care.*

Related CLSI/NCCLS Publications*

C24-A2 **Statistical Quality Control for Quantitative Measurements: Principles and Definitions; Approved Guideline—Second Edition (1999).** This guideline provides definitions of analytical intervals; plans for quality control procedures; and guidance for quality control applications.

C28-A2 **How to Define and Determine Reference Intervals in the Clinical Laboratory; Approved Guideline—Second Edition (2000).** This document provides guidance for determining reference values and reference intervals for quantitative clinical laboratory tests.

EP5-A2 **Evaluation of Precision Performance of Quantitative Measurement Methods; Approved Guideline—Second Edition (2004).** This document provides guidance for designing an experiment to evaluate the precision performance of quantitative measurement methods; recommendations on comparing the resulting precision estimates with manufacturers' precision performance claims and determining when such comparisons are valid; and manufacturers' guidelines for establishing claims.

EP6-A **Evaluation of the Linearity of Quantitative Measurement Procedures: A Statistical Approach; Approved Guideline (2003).** This document provides guidance for characterizing the linearity of a method during a method evaluation; for checking linearity as part of routine quality assurance; and for determining and stating a manufacturer's claim for linear range.

EP7-A **Interference Testing in Clinical Chemistry; Approved Guideline (2002).** This document provides background information, guidance, and experimental procedures for investigating, identifying, and characterizing the effects of interfering substances on clinical chemistry test results.

EP9-A2 **Method Comparison and Bias Estimation Using Patient Samples; Approved Guideline—Second Edition (2002).** This document addresses procedures for determining the bias between two clinical methods or devices, and for the design of a method comparison experiment using split patient samples and data analysis.

EP10-A2 **Preliminary Evaluation of Quantitative Clinical Laboratory Methods; Approved Guideline—Second Edition (2002).** This guideline addresses experimental design and data analysis for preliminary evaluation of the performance of a measurement procedure or device.

GP10-A **Assessment of the Clinical Accuracy of Laboratory Tests Using Receiver Operating Characteristic (ROC) Plots; Approved Guideline (1995).** This document describes the design of a study to evaluate clinical accuracy of laboratory tests and contains procedures for preparing ROC curves; glossary of terms; and information on computer software programs.

* Proposed-level documents are being advanced through the Clinical and Laboratory Standards Institute consensus process; therefore, readers should refer to the most recent editions.

NOTES

NOTES

Active Membership
(as of 1 April 2005)

Sustaining Members

Abbott Laboratories
American Association for Clinical Chemistry
Bayer Corporation
BD
Beckman Coulter, Inc.
bioMérieux, Inc.
CLMA
College of American Pathologists
GlaxoSmithKline
Ortho-Clinical Diagnostics, Inc.
Pfizer Inc
Roche Diagnostics, Inc.

Professional Members

American Academy of Family Physicians
American Association for Clinical Chemistry
American Association for Respiratory Care
American Chemical Society
American Medical Technologists
American Society for Clinical Laboratory Science
American Society for Microbiology
American Society of Hematology
American Type Culture Collection, Inc.
Asociacion Mexicana de Bioquimica Clinica A.C.
Assn. of Public Health Laboratories
Assoc. Micro. Clinici Italiani-A.M.C.L.I.
British Society for Antimicrobial Chemotherapy
Canadian Society for Medical Laboratory Science - Société Canadienne de Science de Laboratoire Médical
Canadian Standards Association
Clinical Laboratory Management Association
COLA
College of American Pathologists
College of Medical Laboratory Technologists of Ontario
College of Physicians and Surgeons of Saskatchewan
ESCMID
International Council for Standardization in Haematology
International Federation of Biomedical Laboratory Science
International Federation of Clinical Chemistry
Italian Society of Clinical Biochemistry and Clinical Molecular Biology
Japan Society of Clinical Chemistry
Japanese Committee for Clinical Laboratory Standards
Joint Commission on Accreditation of Healthcare Organizations
National Academy of Clinical Biochemistry
National Association of Testing Authorities - Australia
National Society for Histotechnology, Inc.
New Zealand Association of Phlebotomy
Ontario Medical Association Quality Management Program-Laboratory Service
RCPA Quality Assurance Programs PTY Limited
Sociedad Espanola de Bioquimica Clinica y Patologia Molecular
Sociedade Brasileira de Analises Clinicas
Taiwanese Committee for Clinical Laboratory Standards (TCCLS)
Turkish Society of Microbiology

Government Members

Armed Forces Institute of Pathology
Association of Public Health Laboratories
BC Centre for Disease Control
Caribbean Epidemiology Centre
Centers for Disease Control and Prevention
Centers for Medicare & Medicaid Services
Centers for Medicare & Medicaid Services/CLIA Program
Chinese Committee for Clinical Laboratory Standards
Commonwealth of Pennsylvania Bureau of Laboratories
Department of Veterans Affairs
Deutsches Institut für Normung (DIN)
FDA Center for Devices and Radiological Health
FDA Center for Veterinary Medicine
FDA Division of Anti-Infective Drug Products
Iowa State Hygienic Laboratory
Massachusetts Department of Public Health Laboratories
National Center of Infectious and Parasitic Diseases (Bulgaria)
National Health Laboratory Service (South Africa)
National Institute of Standards and Technology
National Pathology Accreditation Advisory Council (Australia)
New York State Department of Health
Ontario Ministry of Health
Pennsylvania Dept. of Health
Saskatchewan Health-Provincial Laboratory
Scientific Institute of Public Health; Belgium Ministry of Social Affairs, Public Health and the Environment

Industry Members

AB Biodisk
Abbott Diabetes Care
Abbott Laboratories
Acrometrix Corporation
Advancis Pharmaceutical Corporation
Affymetrix, Inc.
Ammirati Regulatory Consulting
Anna Longwell, PC
A/S ROSCO
AstraZeneca Pharmaceuticals
Aventis
Axis-Shield POC AS
Bayer Corporation - Elkhart, IN
Bayer Corporation - Tarrytown, NY
Bayer Corporation - West Haven, CT
BD
BD Diabetes Care
BD Diagnostic Systems
BD VACUTAINER Systems
Beckman Coulter, Inc.
Beckman Coulter K.K. (Japan)
Bio-Development SRL
Bio-Inova Life Sciences International
Biomedia Laboratories SDN BHD
bioMérieux, Inc. (MO)
Biometrology Consultants
Bio-Rad Laboratories, Inc.
Bio-Rad Laboratories, Inc. – France
Bio-Rad Laboratories, Inc. – Plano, TX
Blaine Healthcare Associates, Inc.
Bristol-Myers Squibb Company
Canadian External Quality Assessment Laboratory
Cepheid
Chen & Chen, LLC
Chiron Corporation
ChromaVision Medical Systems, Inc.
Clinical Micro Sensors
The Clinical Microbiology Institute
Cognigen
CONOSCO
Copan Diagnostics Inc.
Cosmetic Ingredient Review
Cubist Pharmaceuticals
Dade Behring Inc. - Cupertino, CA
Dade Behring Inc. - Deerfield, IL
Dade Behring Inc. - Glasgow, DE
Dade Behring Inc. - Marburg, Germany
Dade Behring Inc. - Sacramento, CA
David G. Rhoads Associates, Inc.
Diagnostic Products Corporation
Digene Corporation
Eiken Chemical Company, Ltd.
Elanco Animal Health
Electa Lab s.r.l.
Enterprise Analysis Corporation
F. Hoffman-La Roche AG
Gen-Probe
GlaxoSmithKline
Greiner Bio-One Inc.
Immunicon Corporation
ImmunoSite, Inc.
Instrumentation Laboratory
International Technidyne Corporation
I-STAT Corporation
Johnson and Johnson Pharmaceutical Research and Development, L.L.C.
K.C.J. Enterprises
LabNow, Inc.
LifeScan, Inc. (a Johnson & Johnson Company)
Machaon Diagnostics
Medical Device Consultants, Inc.
Merck & Company, Inc.
Micromyx, LLC
Minigrip/Zip-Pak
Nanosphere, Inc.
National Pathology Accreditation Advisory Council (Australia)
Nippon Becton Dickinson Co., Ltd.
Nissui Pharmaceutical Co., Ltd.
Novartis Pharmaceuticals Corporation
Olympus America, Inc.
Optimer Pharmaceuticals, Inc.
Ortho-Clinical Diagnostics, Inc. (Rochester, NY)
Ortho-McNeil Pharmaceutical (Raritan, NJ)
Oxoid Inc.
Paratek Pharmaceuticals
Pfizer Animal Health
Pfizer Inc
Pfizer Italia Srl
Powers Consulting Services
Predicant Biosciences
Procter & Gamble Pharmaceuticals, Inc.
QSE Consulting
Radiometer America, Inc.
Radiometer Medical A/S
Replidyne
Roche Diagnostics GmbH
Roche Diagnostics, Inc.
Roche Diagnostics Shanghai Ltd.
Roche Laboratories (Div. Hoffmann-La Roche Inc.)
Sanofi Pasteur
Sarstedt, Inc.
Schering Corporation
Schleicher & Schuell, Inc.
SFBC Anapharm
Streck Laboratories, Inc.
SYN X Pharma Inc.
Sysmex Corporation (Japan)
Sysmex Corporation (Long Grove, IL)
TheraDoc
Theravance Inc.
Thrombodyne, Inc.
THYMED GmbH
Transasia Engineers
Trek Diagnostic Systems, Inc.
Vetoquinol S.A.
Vicuron Pharmaceuticals Inc.
Vysis, Inc.
Wyeth Research
XDX, Inc.
YD Consultant
YD Diagnostics (Seoul, Korea)

Trade Associations

AdvaMed
Japan Association of Clinical Reagents Industries (Tokyo, Japan)

Associate Active Members

82 MDG/SGSCL (Sheppard AFB,TX)
Academisch Ziekenhuis -VUB (Belgium)
ACL Laboratories (WI)
All Children's Hospital (FL)
Allegheny General Hospital (PA)
Allina Health System (MN)
American University of Beirut Medical Center (NY)
Anne Arundel Medical Center (MD)
Antwerp University Hospital (Belgium)
Arkansas Department of Health
ARUP at University Hospital (UT)
Associated Regional & University Pathologists (UT)
Atlantic Health System (NJ)
AZ Sint-Jan (Belgium)
Azienda Ospedale Di Lecco (Italy)
Barnes-Jewish Hospital (MO)
Baxter Regional Medical Center (AR)
Baystate Medical Center (MA)
Bbaguas Duzen Laboratories (Turkey)
BC Biomedical Laboratories (Surrey, BC, Canada)
Bermuda Hospitals Board
Bo Ali Hospital (Iran)
Bon Secours Hospital (Ireland)
Brazosport Memorial Hospital (TX)
Broward General Medical Center (FL)
Cadham Provincial Laboratory (Winnipeg, MB, Canada)
Calgary Laboratory Services (Calgary, AB, Canada)
California Pacific Medical Center
Cambridge Memorial Hospital (Cambridge, ON, Canada)
Canterbury Health Laboratories (New Zealand)
Cape Breton Healthcare Complex (Nova Scotia, Canada)
Carilion Consolidated Laboratory (VA)
Carolinas Medical Center (NC)
Cathay General Hospital (Taiwan)
Central Laboratory for Veterinarians (BC, Canada)
Central Ohio Primary Care Physicians
Central Texas Veterans Health Care System
Centro Diagnostico Italiano (Milano, Italy)
Chang Gung Memorial Hospital (Taiwan)
Changi General Hospital (Singapore)
Children's Hospital (NE)
Children's Hospital Central California
Children's Hospital & Clinics (MN)
Children's Hospital Medical Center (Akron, OH)
Children's Medical Center of Dallas (TX)
Chinese Association of Advanced Blood Bankers (Beijing)
CHR St. Joseph Warquignies (Belgium)
City of Hope National Medical Center (CA)
Clarian Health - Methodist Hospital (IN)
CLSI Laboratories (PA)
Community Hospital of Lancaster (PA)
Community Hospital of the Monterey Peninsula (CA)
CompuNet Clinical Laboratories (OH)
Covance Central Laboratory Services (IN)
Creighton University Medical Center (NE)
Detroit Health Department (MI)
DFS/CLIA Certification (NC)
Diagnostic Accreditation Program (Vancouver, BC, Canada)
Diagnósticos da América S/A (Brazil)
Dianon Systems (OK)
Dr. Everett Chalmers Hospital (New Brunswick, Canada)
Duke University Medical Center (NC)
Dwight David Eisenhower Army Medical Center (GA)
Eastern Health Pathology (Australia)
Emory University Hospital (GA)
Enzo Clinical Labs (NY)
Evangelical Community Hospital (PA)
Fairview-University Medical Center (MN)
Florida Hospital East Orlando
Focus Technologies (CA)
Focus Technologies (VA)
Foothills Hospital (Calgary, AB, Canada)
Franciscan Shared Laboratory (WI)
Fresno Community Hospital and Medical Center
Gamma Dynacare Medical Laboratories (Ontario, Canada)
Gateway Medical Center (TN)
Geisinger Medical Center (PA)
Guthrie Clinic Laboratories (PA)
Hagerstown Medical Laboratory (MD)
Harris Methodist Fort Worth (TX)
Hartford Hospital (CT)
Headwaters Health Authority (Alberta, Canada)

Health Network Lab (PA)
Highlands Regional Medical Center (FL)
Hoag Memorial Hospital Presbyterian (CA)
Holy Cross Hospital (MD)
Hôpital Maisonneuve - Rosemont (Montreal, Canada)
Hôpital Saint-Luc (Montreal, Quebec, Canada)
Hospital Consolidated Laboratories (MI)
Hospital de Sousa Martins (Portugal)
Hospital for Sick Children (Toronto, ON, Canada)
Hotel Dieu Grace Hospital (Windsor, ON, Canada)
Huddinge University Hospital (Sweden)
Humility of Mary Health Partners (OH)
Hunter Area Health Service (Australia)
Hunterdon Medical Center (NJ)
Indiana University
Innova Fairfax Hospital (VA)
Institute of Medical and Veterinary Science (Australia)
International Health Management Associates, Inc. (IL)
Jackson Health System (FL)
Jacobi Medical Center (NY)
John H. Stroger, Jr. Hospital of Cook County (IL)
Johns Hopkins Medical Institutions (MD)
Kadlec Medical Center (WA)
Kaiser Permanente (MD)
Kantonsspital (Switzerland)
Kimball Medical Center (NJ)
King Abdulaziz Medical City – Jeddah (Saudi Arabia)
King Faisal Specialist Hospital (Saudi Arabia)
LabCorp (NC)
Laboratoire de Santé Publique du Quebec (Canada)
Laboratorio Dr. Echevarne (Spain)
Laboratório Fleury S/C Ltda. (Brazil)
Laboratorio Manlab (Argentina)
Laboratory Corporation of America (NJ)
Lakeland Regional Medical Center (FL)
Landstuhl Regional Medical Center (APO AE)
Lawrence General Hospital (MA)
Lewis-Gale Medical Center (VA)
L'Hotel-Dieu de Quebec (Canada)
Libero Instituto Univ. Campus BioMedico (Italy)
Lindy Boggs Medical Center (LA)
Loma Linda Mercantile (CA)
Long Beach Memorial Medical Center (CA)
Los Angeles County Public Health Lab (CA)
Lourdes Hospital (KY)
Maimonides Medical Center (NY)
Marion County Health Department (IN)
Martin Luther King/Drew Medical Center (CA)
Massachusetts General Hospital (Microbiology Laboratory)
MDS Metro Laboratory Services (Burnaby, BC, Canada)
Medical College of Virginia Hospital
Medical Research Laboratories International (KY)
Medical University of South Carolina
Memorial Medical Center (Napoleon Avenue, New Orleans, LA)
Methodist Hospital (Houston, TX)
Methodist Hospital (San Antonio, TX)
Mid America Clinical Laboratories, LLC (IN)
Middlesex Hospital (CT)
Montreal Children's Hospital (Canada)
Montreal General Hospital (Canada)
National Serology Reference Laboratory (Australia)
NB Department of Health & Wellness (New Brunswick, Canada)
The Nebraska Medical Center
Nevada Cancer Institute
New Britain General Hospital (CT)
New England Fertility Institute (CT)
New York City Department of Health & Mental Hygiene
NorDx (ME)
North Carolina State Laboratory of Public Health
North Central Medical Center (TX)
North Shore - Long Island Jewish Health System Laboratories (NY)
North Shore University Hospital (NY)
Northwestern Memorial Hospital (IL)
Ochsner Clinic Foundation (LA)
Onze Lieve Vrouw Ziekenhuis (Belgium)
Orlando Regional Healthcare System (FL)
Ospedali Riuniti (Italy)
The Ottawa Hospital (Ottawa, ON, Canada)
Our Lady of the Resurrection Medical Center (IL)
Pathology and Cytology Laboratories, Inc. (KY)
Pathology Associates Medical Laboratories (WA)
The Permanente Medical Group (CA)
Phoenix College (AZ)
Piedmont Hospital (GA)
Pocono Medical Center (PA)
Presbyterian Hospital of Dallas (TX)
Providence Health Care (Vancouver, BC, Canada)
Provincial Laboratory for Public Health (Edmonton, AB, Canada)
Quest Diagnostics Incorporated (CA)
Quintiles Laboratories, Ltd. (GA)
Regional Health Authority Four (NB, Canada)
Regions Hospital
Rex Healthcare (NC)
Rhode Island Department of Health Laboratories
Riverside Medical Center (IL)
Robert Wood Johnson University Hospital (NJ)
Sahlgrenska Universitetssjukhuset (Sweden)
St. Alexius Medical Center (ND)
St. Anthony Hospital (CO)
St. Anthony's Hospital (FL)
St. Barnabas Medical Center (NJ)
St. Christopher's Hospital for Children (PA)
St-Eustache Hospital (Quebec, Canada)
St. John Hospital and Medical Center (MI)
St. John Regional Hospital (St. John, NB, Canada)
St. John's Hospital & Health Center (CA)
St. Joseph's Hospital – Marshfield Clinic (WI)
St. Jude Children's Research Hospital (TN)
St. Mary Medical Center (CA)
St. Mary of the Plains Hospital (TX)
St. Michael's Hospital (Toronto, ON, Canada)
St. Vincent's University Hospital (Ireland)
Ste. Justine Hospital (Montreal, PQ, Canada)
Salem Clinic (OR)
San Francisco General Hospital (CA)
Santa Clara Valley Medical Center (CA)
Seoul Nat'l University Hospital (Korea)
Shands at the University of Florida
South Bend Medical Foundation (IN)
South Western Area Pathology Service (Australia)
Southern Maine Medical Center
Spartanburg Regional Medical Center (SC)
Specialty Laboratories, Inc. (CA)
State of Connecticut Dept. of Public Health
State of Washington Department of Health
Stony Brook University Hospital (NY)
Stormont-Vail Regional Medical Center (KS)
Sun Health-Boswell Hospital (AZ)
Sunnybrook Health Science Center (ON, Canada)
Sunrise Hospital and Medical Center (NV)
Swedish Medical Center - Providence Campus (WA)
Temple University Hospital (PA)
Tenet Odessa Regional Hospital (TX)
Touro Infirmary (LA)
Tripler Army Medical Center (HI)
Truman Medical Center (MO)
Tuen Mun Hospital (Hong Kong)
UCLA Medical Center (CA)
UCSF Medical Center (CA)
UNC Hospitals (NC)
Unidad de Patologia Clinica (Mexico)
Union Clinical Laboratory (Taiwan)
United Laboratories Company (Kuwait)
Universita Campus Bio-Medico (Italy)
University College Hospital (Galway, Ireland)
University of Chicago Hospitals (IL)
University of Colorado Hospital
University of Debrecen Medical Health and Science Center (Hungary)
University of Illinois Medical Center
University of Maryland Medical System
University of Medicine & Dentistry, NJ University Hospital
University of the Ryukyus (Japan)
University of Wisconsin Hospital
The University of the West Indies
University of Virginia Medical Center
University of Washington
US LABS, Inc. (CA)
USA MEDDAC-AK
UZ-KUL Medical Center (Belgium)
VA (Tuskegee) Medical Center (AL)
Virginia Beach General Hospital (VA)
Virginia Department of Health
Washington Adventist Hospital (MD)
Washoe Medical Center Laboratory (NV)
Waterford Regional Hospital (Ireland)
Wellstar Health Systems (GA)
West China Second University Hospital, Sichuan University (P.R. China)
West Jefferson Medical Center (LA)
Wilford Hall Medical Center (TX)
William Beaumont Army Medical Center (TX)
William Beaumont Hospital (MI)
Winn Army Community Hospital (GA)
Winnipeg Regional Health Authority (Winnipeg, Canada)
Wishard Memorial Hospital (IN)
York Hospital (PA)

940 West Valley Road ▾ Suite 1400 ▾ Wayne, PA 19087 ▾ USA ▾ PHONE 610.688.0100
FAX 610.688.0700 ▾ E-MAIL: customerservice@clsi.org ▾ WEBSITE: www.clsi.org ▾ ISBN 1-56238-574-7

CLINICAL AND LABORATORY STANDARDS INSTITUTE ™

(Formerly NCCLS)
Providing NCCLS standards and guidelines,
ISO/TC 212 standards, and ISO/TC 76 standards