CLINICAL AND LABORATORY STANDARDS INSTITUTE®

August 2013

EP09-A3

Measurement Procedure Comparison and Bias Estimation Using Patient Samples; Approved Guideline—Third Edition

This document addresses the design of measurement procedure comparison experiments using patient samples and subsequent data analysis techniques used to determine the bias between two *in vitro* diagnostic measurement procedures.

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

Clinical and Laboratory Standards Institute

Setting the standard for quality in clinical laboratory testing around the world.

The Clinical and Laboratory Standards Institute (CLSI) is a not-for-profit membership organization that brings together the varied perspectives and expertise of the worldwide laboratory community for the advancement of a common cause: to foster excellence in laboratory medicine by developing and implementing clinical laboratory standards and guidelines that help laboratories fulfill their responsibilities with efficiency, effectiveness, and global applicability.

Consensus Process

Consensus—the substantial agreement by materially affected, competent, and interested parties—is core to the development of all CLSI documents. It does not always connote unanimous agreement, but does mean that the participants in the development of a consensus document have considered and resolved all relevant objections and accept the resulting agreement.

Commenting on Documents

CLSI documents undergo periodic evaluation and modification to keep pace with advancements in technologies, procedures, methods, and protocols affecting the laboratory or health care.

CLSI's consensus process depends on experts who volunteer to serve as contributing authors and/or as participants in the reviewing and commenting process. At the end of each comment period, the committee that developed the document is obligated to review all comments, respond in writing to all substantive comments, and revise the draft document as appropriate.

Comments on published CLSI documents are equally essential, and may be submitted by anyone, at any time, on any document. All comments are addressed according to the consensus process by a committee of experts.

Appeals Process

If it is believed that an objection has not been adequately addressed, the process for appeals is documented in the CLSI Standards Development Policies and Process document.

All comments and responses submitted on draft and published documents are retained on file at CLSI and are available upon request.

Get Involved—Volunteer!

Do you use CLSI documents in your workplace? Do you see room for improvement? Would you like to get involved in the revision process? Or maybe you see a need to develop a new document for an emerging technology? CLSI wants to hear from you. We are always looking for volunteers. By donating your time and talents to improve the standards that affect your own work, you will play an active role in improving public health across the globe.

For further information on committee participation or to submit comments, contact CLSI.

Clinical and Laboratory Standards Institute
950 West Valley Road, Suite 2500
Wayne, PA 19087 USA
P: 610.688.0100
F: 610.688.0700
www.clsi.org
standard@clsi.org

ISBN 1-56238-887-8 (Print)
ISBN 1-56238-888-6 (Electronic)
ISSN 1558-6502 (Print)
ISSN 2162-2914 (Electronic)

EP09-A3
Vol. 33 No. 11
Replaces EP09-A2-IR
Vol. 30 No. 17

Measurement Procedure Comparison and Bias Estimation Using Patient Samples; Approved Guideline—Third Edition

Volume 33 Number 11

Jeffrey R. Budd, PhD
A. Paul Durham, MA
Thomas E. Gwise, PhD
Beimar Iriarte, MS
Anders Kallner, MD, PhD
Kristian Linnet, MD, PhD
Robert Magari, PhD
Jeffrey E. Vaks, PhD

Abstract

Clinical and Laboratory Standards Institute document EP09-A3—*Measurement Procedure Comparison and Bias Estimation Using Patient Samples; Approved Guideline—Third Edition* is written for laboratorians and manufacturers. It describes procedures for determining the bias between two measurement procedures, and it identifies factors for consideration when designing and analyzing a measurement procedure comparison experiment using split patient samples. An overview of the measurement procedure comparison experiment includes considerations for both manufacturers and laboratorians. Details on how to create difference and scatter plots for visual inspection of the data are provided. Once the data are characterized, various methods are introduced for quantifying the relationship between two measurement procedures, including bias estimates and regression techniques. The final section contains recommendations for manufacturers' evaluation of bias and statement format for bias claims.

Clinical and Laboratory Standards Institute (CLSI). *Measurement Procedure Comparison and Bias Estimation Using Patient Samples; Approved Guideline—Third Edition*. CLSI document EP09-A3 (ISBN 1-56238-887-8 [Print]; ISBN 1-56238-888-6 [Electronic]). Clinical and Laboratory Standards Institute, 950 West Valley Road, Suite 2500, Wayne, Pennsylvania 19087 USA, 2013.

Suggested Citation

CLSI. *Measurement Procedure Comparison and Bias Estimation Using Patient Samples; Approved Guideline—Third Edition.* CLSI document EP09-A3. Wayne, PA: Clinical and Laboratory Standards Institute; 2013.

Proposed Guideline
January 1986

Tentative Guideline
April 1993

Approved Guideline
December 1995

Approved Guideline—Second Edition
September 2002

Approved Guideline—Second Edition (Interim Revision)
July 2010

Approved Guideline—Third Edition
August 2013

ISBN 1-56238-887-8 (Print)
ISBN 1-56238-888-6 (Electronic)
ISSN 1558-6502 (Print)
ISSN 2162-2914 (Electronic)

Committee Membership

Consensus Committee on Evaluation Protocols

James F. Pierson-Perry
Chairholder
Siemens Healthcare Diagnostics
Newark, Delaware, USA

Mitchell G. Scott, PhD
Vice-Chairholder
Barnes-Jewish Hospital
Washington University
School of Medicine
St. Louis, Missouri, USA

Rex Astles, PhD, FACB, DABCC
Centers for Disease Control and Prevention
Atlanta, Georgia, USA

Jeffrey R. Budd, PhD
Beckman Coulter
Chaska, Minnesota, USA

Karl De Vore
Bio-Rad Laboratories, Inc.
Irvine, California, USA

Jonathan Guy Middle, PhD
University Hospital Birmingham NHS Trust
Birmingham, United Kingdom

James H. Nichols, PhD, DABCC, FACB
Vanderbilt University
School of Medicine
Nashville, Tennessee, USA

Gene Pennello, PhD
FDA Center for Devices and Radiological Health
Silver Spring, Maryland, USA

Document Development Committee on Method Comparison and Bias Estimation Using Patient Samples

Jeffrey R. Budd, PhD
Chairholder
Beckman Coulter
Chaska, Minnesota, USA

A. Paul Durham, MA
Culver City, California, USA

Thomas E. Gwise, PhD
FDA Center for Drug Evaluation and Research
Silver Spring, Maryland, USA

Anders Kallner, MD, PhD
Karolinska Hospital
Stockholm, Sweden

Kristian Linnet, MD, PhD
University of Copenhagen
Copenhagen, Denmark

Staff

Clinical and Laboratory Standards Institute
Wayne, Pennsylvania, USA

Luann Ochs, MS
Senior Vice President – Operations

Ron S. Quicho
Staff Liaison

Patrice E. Polgar
Project Manager

Megan L. Tertel, MA
Editor

Acknowledgment

CLSI and the Consensus Committee on Evaluation Protocols gratefully acknowledge the following volunteers for their important contributions to the development of this document:

Beimar Iriarte, MS
Abbott Diagnostics
Abbott Park, Illinois, USA

Robert Magari, PhD
Beckman Coulter
Miami, Florida, USA

Jeffrey E. Vaks, PhD
Roche Molecular Diagnostics
Pleasanton, California, USA

Contents

Contents (Continued)

Foreword

Measurement procedure comparison is one of the most common techniques used by both manufacturers and clinical laboratorians to estimate the bias of an *in vitro* diagnostic (IVD) measurement procedure relative to a comparator. It involves the comparison of results from patient samples from two measurement procedures intended to measure the same component (eg, concentration of a measurand) with the key determination being the estimate of bias between them.

A number of different scenarios exist in which measurement procedure comparison studies are indicated. For both the manufacturer and the clinical laboratorian, the ideal scenario is the comparison of a candidate measurement procedure to a generally accepted standard or reference measurement procedure. In the case of a manufacturer, this involves the establishment and perhaps verification of performance claims for bias, while in the case of a laboratorian, it involves introducing a measurement procedure into the laboratory, including verification of such manufacturer claims (specifications). The scope of the experimental and data-handling procedures for these two purposes will differ. In either case the assumption that the reference measurement procedure provides "true" values means that bias (systematic measurement error) is estimated.

Quite commonly, however, there is no standard or reference measurement procedure. The manufacturer instead compares a candidate measurement procedure to the best measurement procedure currently available. The laboratorian usually compares the candidate and an available procedure. Then, there may not be a "true" value and the "difference," rather than the "bias," is estimated.

Given the variety of performance characteristics of IVD measurement procedures, a single experimental design is not appropriate for all types of laboratorian and manufacturer measurement procedure comparisons. Therefore, performance characteristics such as measuring interval and precision profile are taken into account in structuring an experiment for comparing two measurement procedures. Multiple worked examples are presented.

This document is intended to promote effective and correct data analysis and reporting using standard experimental and statistical methods.

It is recommended that manufacturers of clinical laboratory measurement procedures and/or devices use this document to establish and standardize their bias performance claims. Many different forms have been used for such claims, and they have not always been sufficiently specific to allow user verification.

A number of changes and additions are included in this revision of the document, including:

- Broader coverage of method comparison applications
- More reasons for comparisons based on patient samples (factor comparisons [eg, sample tube types])
- Visualization/exploration of data using difference plots
- Regression descriptions including weighted options, Deming, and Passing-Bablok techniques
- Measurement of bias using difference plots
- Measurement of bias at clinical decision points
- Computation of confidence intervals for all parameters

- Outlier detection using extreme studentized deviate
- Relocation of most of the detailed mathematical descriptions to the appendixes

> **NOTE:** Due to the complex nature of the calculations in this guideline, it is recommended that the user have access to a computer and statistical software, such as StatisPro™ method evaluation software from CLSI.

Key Words

Alternative regression methods, bias, evaluation protocol, experimental design, linear regression, measurement procedure comparison, outliers, quality control, residuals

Measurement Procedure Comparison and Bias Estimation Using Patient Samples; Approved Guideline—Third Edition

1 Scope

This document provides guidance for designing an experiment and selecting methods to quantify systematic measurement error (bias or difference) between measurement procedures based on comparing patient samples. It provides procedures to determine the average bias between two measurement procedures either across their measuring intervals or at selected concentrations. Intended users of this guideline are manufacturers of *in vitro* diagnostic (IVD) reagents—which includes those who create laboratory-developed tests—as well as regulatory bodies and clinical laboratory personnel.

This document is for use with measurement procedures that provide quantitative numerical results. This document is not intended for use with ordinal IVD measurement procedures, commonly referred to as qualitative procedures (see CLSI document EP12[1]). This document is not intended to address evaluation of random error (see CLSI documents EP05[2] and EP15[3]) or to determine the total error inherent in a comparison of measurement procedures (see CLSI document EP21[4]). It is not intended to measure the variability of multiple replicates collected during the measurement of a sample, nor is it intended to measure the bias of individual measurements such as those resulting from sample interference (as covered in CLSI document EP07[5]).

2 Introduction

The purpose of this document is to establish good practices at measuring average bias over the measuring interval in a population of patient samples, relative to a comparative or reference method. Difference plots are used to visually portray the relationship between measurement procedures to evaluate if the relationship is consistent with a constant difference or as a constant percentage difference (constant CV) over the measuring interval. The plots are also used to determine the bias estimate from such plots through either an average or a median. Given the knowledge gained from the difference plots, users are provided with regression fit options to characterize bias in terms of slope and intercept and bias estimates at selected concentrations.

This document describes multiple situations in which measurement procedures are compared, each of which has its own experimental requirements. These requirements dictate differences in the number of factors to incorporate into the experimental design, the number of samples, and the number of replicates for each sample. The situations covered in this document assume a study is comparing two procedures that measure the same quantity by using measurement procedure results from study samples.

In selecting an analysis technique for a set of data, a stepwise process is described that starts with visual data inspection using difference and scatter plots. The data from difference plots can then be used to estimate the bias (or percent bias) between measurement procedures. Clinical laboratorians may require no further analysis. The document continues, however, by describing various regression techniques and their underlying assumptions that help determine which one should be used in each situation. Such techniques can, in many cases, provide more robust estimates of bias, so clinical laboratories may wish to use them. Manufacturers will use them in almost all cases. The goal throughout the document is to propose a set of techniques for determining bias between measurement procedures and to detail the strengths and weaknesses of these techniques for given situations.

A brief description of measurement procedure comparison scenarios is provided in the following sections. Section 2.1 is a general overview common to all scenarios. Sections 2.2.1 and 2.2.2 are intended for

manufacturers or research laboratories that have created a candidate measurement procedure. Section 2.2.3 is intended for the typical clinical laboratorian.

2.1 Overview of the Measurement Procedure Comparison Study

The purpose of a measurement procedure comparison study must first be determined, along with the two measurement procedures to be used in the study. The three primary purposes for such a study are introduced in Section 2.2. Other purposes for such studies are described in Section 10.

For any well-conducted study, the personnel performing the measurements must be familiar with the instrument systems used in the study. The familiarization period is described in Section 5.

There are numerous considerations for conducting any measurement procedure comparison study. General considerations, including sample selection and handling, are covered in Section 6, while those specific only for studies by clinical laboratories are covered in Section 7.

Once the data are collected for the study, they must be reviewed to determine if the goals of data collection have been met and to characterize the data for interval and distribution of measurements, and other factors that can dictate what techniques are used for data analysis. Visual data inspection techniques are suggested in Section 8 for this purpose.

The outcome of a measurement procedure comparison study is a quantification of the bias between two such procedures. This bias can be expressed as an average bias over the measured interval or a bias at a selected concentration. The techniques that can be used to supply such bias estimates are described in Section 9.

Finally, Section 11 discusses the steps for comparing estimated bias to acceptance criteria and for stating performance claims.

Throughout this document, the terms in Table 1 are used to describe the measurement procedures to be compared.

Table 1. Measurement Procedure Terminology

X **Characteristic Plotted on the Horizontal, x-axis**	***Y*** **Characteristic Plotted on the Vertical, y-axis**
Independent variable	Dependent variable
Comparative measurement procedure	Candidate measurement procedure
Reference measurement procedure	

2.2 Primary Purposes for Measurement Procedure Comparisons

2.2.1 Establishing the Relationship Between Measurement Procedures by the Manufacturer

Manufacturers must establish the relationship of any candidate measurement procedure of measurand quantification with a comparative measurement procedure, ideally a reference measurement procedure. Typically, when such a reference measurement procedure is available, the desired result of the comparison is no significant bias between them. However, often a new (candidate) measurement procedure is developed as an improvement over a comparative measurement procedure (eg, an automated *in vitro* diagnostic (IVD) procedure to replace a microplate procedure). In such situations the primary goal is to establish the bias between them. It is recommended that at least 100 patient samples with measurand values spanning the common measuring interval of the two measurement procedures be used for establishment of bias claims. Influential factors can be included in the experimental design such as

calibration, run, day, reagent lot, calibrator lot, and instrument. The average result of multiple sample replicates may be used for both procedures to decrease the uncertainty of the bias estimate.

2.2.2 Claims Verification by the Manufacturer (Validation)

The second measurement procedure comparison situation a manufacturer may encounter is the verification that an IVD measurement procedure meets the claims already established for it. Such a verification, more than any other type of measurement procedure performance study, can be the key to validating that the measurement procedure is fit for the purpose of quantitatively determining the concentration of the measurand in a clinical laboratory setting. Therefore, such an experiment is usually performed at one or more external sites, because the goal is to show that the candidate measurement procedure has low bias in comparison to a trusted comparative measurement procedure in an in-use situation.

Generally speaking, to validate that a candidate measurement procedure is fit for purpose, it is evaluated per its instructions for use. Usually, the procedure uses only a single replicate. Statistically, this fact does not invalidate the estimation of bias using averages of multiple replicates. Provided each replicate represents equivalent information (ie, order has no influence), each represents a result from the procedure per instructions for use, so averaging of the results on multiple replicates simply improves bias estimation. Multiple replicates also permit Deming regression (see Appendixes E and F) without requiring prior knowledge of measurement procedure imprecision. For the visual displays of relative bias presented in Sections 8 and 9, replicate results from each measuring procedure (candidate, comparator) would be averaged before being plotted.

However, given that the primary purpose of this type of verification study is to validate whether the candidate measurement procedure is fit for purpose, it may be required 1) to use a set of single candidate results from individual replicates matched to average comparator results across multiple replicates or 2) to use a set of single candidate results matched with single comparator results. In these two cases, the visual displays described in Sections 8 and 9 may be from individual measurements rather than averaged results. Analysis techniques remain the same, except the use of Deming regression requires prior knowledge of measurement procedure imprecision.

It is recommended that at least 100 patient samples with measurand values spanning as much of the common measuring interval of both measurement procedures as feasible be used for such a validation, and that the study be conducted at each site over three to five days at minimum. Typically, the measurement procedure's bias claims are verified if the estimated bias is within a predetermined acceptance criterion.

2.2.3 Measurement Procedure Introduction to the Clinical Laboratory

Clinical laboratories typically perform measurement procedure comparison studies when they are introducing an IVD product into their menu. The candidate measurement procedure typically replaces one currently used in the laboratory. A decision to bring in a candidate measurement procedure is often based, at least in part, on the performance results provided by the manufacturer via either performance claims or postmarket study comparisons. In either case, if the comparison to be performed by the laboratory is available from the manufacturer, the desired outcome is to verify the manufacturer-supplied bias performance. Otherwise the goal is to independently quantify the bias (difference). The bias can help determine what changes, if any, need to be made in reporting results from the candidate measurement procedure. This may include changes in reference intervals (see CLSI document EP28[6]) or medical decision values to reflect the difference between measurement procedure results.

To perform the analysis methods described in this guideline, clinical laboratories should attempt to measure at least 40 patient samples that span the measuring interval of the measurement procedures.

Single, nonreplicated sample measurements are typically used within each measurement procedure. If sample volume and time restraints permit, the average result of multiple sample replicates may be used for both procedures to decrease the uncertainty of the bias estimate.

2.2.4 Summary of Measurement Procedure Comparison Studies

Table 2 lists the characteristics typical of each type of study.

Table 2. Typical Study Characteristics

Type of Study	Conducted by	Number of Samples	Number of Candidate MP Replicates Used	Number of Candidate MP Systems	Recommended Analyses to Be Used to Determine Bias
MP Claims Establishment	Manufacturer	≥100	1 or more	1 or more	Regression
MP Claims Verification (Validation)	Manufacturer	≥100	1	1 or more	Regression
MP Introduction (Verification)	Laboratory	≥40	1 or more	1	Difference plot or regression

Abbreviation: MP, measurement procedure.

There may or may not be restrictions on the number of replicates for the comparative measurement procedure. Whenever multiple replicates are used, the average of a sample's replicates is typically used as the estimate of the result for each sample on that measurement procedure.

There are no restrictions on the number of candidate measurement systems to be used by a manufacturer. A laboratory will usually have only one candidate measurement system that they are introducing into their laboratory. The goal is usually to determine how that system compares to their current measurement system (procedure).

3 Standard Precautions

Because it is often impossible to know what isolates or specimens 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 known infectious agents and thus are more comprehensive than universal precautions, which are intended to apply only to transmission of blood-borne pathogens. The Centers for Disease Control and Prevention address this topic in published guidelines that address the daily operations of diagnostic medicine in human and animal medicine while encouraging a culture of safety in the laboratory.[7] For specific precautions for preventing the laboratory transmission of all known infectious agents from laboratory instruments and materials and for recommendations for the management of exposure to all known infectious diseases, refer to CLSI document M29.[8]

4 Terminology

4.1 A Note on Terminology

CLSI, as a global leader in standardization, is firmly committed to achieving global harmonization wherever possible. Harmonization is a process of recognizing, understanding, and explaining differences while taking steps to achieve worldwide uniformity. CLSI recognizes 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, International Organization for Standardization (ISO), and European Committee for Standardization (CEN) documents; and that legally required use of terms, regional usage, and different consensus timelines are all important considerations in the harmonization process. In light of this, CLSI's consensus process for development and revision of standards and guidelines focuses on harmonization of terms to facilitate the global application of standards and guidelines.

Essentially, new documents must adhere to the latest edition of the *International vocabulary of metrology — Basic and general concepts and associated terms* (VIM)[9] whenever an ambiguity in the interpretation or understanding of terms occurs. In the latest edition of the VIM, many definitions have become more explicit and understandable, but the language of the VIM is difficult and compact. VIM deals with general metrology and terminology that should be useful for most disciplines that measure quantities.

The understanding of a few terms has changed during the last decade as the concepts have developed. Particularly, *trueness* (measurement trueness) is defined as expressing the closeness of agreement between the average of an infinite number of replicate measurements and a reference value; and *precision* (measurement precision) is defined as closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions. Consequently, *accuracy* (measurement accuracy) is the closeness of agreement between a measured value and a true quantity value of a measurand. Thus, this concept comprises both trueness and precision, and applies to a single result. *Measuring interval* has replaced *reportable range* when referring to "a set of values of a measurand for which the error of a measuring instrument (test) is intended to lie within specified limits." An *interval* [*a*;*b*] is delineated by two limits *a* and *b* ($b > a$), whereas a *range* (r[*a*;*b*]) is expressed as the difference between *b* and *a* ($b - a$). Thus, the range of the interval [*a*;*b*] is the difference ($b - a$) that is denoted by *r[a;b]*.

The term *measurand* is used when referring to the quantity intended to be measured instead of *analyte* (component represented in the name of a measurable quantity). The term *measurement procedure* replaces *analytical method* and *assay* for a set of operations, used in the performance of particular measurements according to a given method.

Verification focuses on whether specifications of a measurement procedure can be achieved, whereas *validation* verifies that the procedure is fit for purpose.

4.2 Definitions

accuracy (measurement) – closeness of agreement between a measured quantity value and a true quantity value of a measurand (JCGM 200:2012)[9]; **NOTE 1:** The concept "measurement accuracy" is not a quantity and is not given a numerical quantity value. A measurement is said to be more accurate when it offers a smaller measurement error (JCGM 200:2012)[9]; **NOTE 2:** The term "measurement accuracy" should not be used for "measurement trueness" and the term "measurement precision" should not be used for "measurement accuracy," which, however, is related to both these concepts (JCGM 200:2012)[9]; **NOTE 3:** "Measurement accuracy" is sometimes understood as closeness of agreement between measured quantity values that are being attributed to the measurand (JCGM 200:2012).[9]

bias (of measurement) – estimate of a systematic measurement error (JCGM 200:2012)[9]; **NOTE 1:** In the context of this guideline, bias refers to the estimated, average bias over the measuring interval from a measurement procedure comparison study; **NOTE 2:** In this document, the metrological term "bias" is equivalent to the term "difference."

correlation coefficient (*r*)//determination coefficient (r^2) – a measure of the linear relationship between two random variables; **NOTE 1:** It ranges from −1 to 1, ie, from perfect negative to perfect positive linear relationship; **NOTE 2:** $r = 0$ indicates no observed linear relationship.

decision point (medical decision point) – a concentration of the measurand that is used as a threshold for making a clinical statement; **NOTE:** Often, decision points will refer to reference limits, but other concentrations, such as from clinical guidelines, are also used.

Deming regression – a method to estimate slope and intercept parameters from a measurement procedure comparison experiment with allowance for both measurement procedures to have imprecision; **NOTE:** The measurement error for each measurement procedure is accounted for in the estimation procedure.[10]

difference plot – a plot of the difference between a measured value and a reference concentration plotted on the y-axis vs the reference concentration on the x-axis; **NOTE 1:** Often, a dashed line is drawn at zero difference; **NOTE 2:** The reference concentration is often expressed as the average of the results of the measurements; **NOTE 3:** The difference may be expressed relative to the reference concentration.

imprecision – the random dispersion of a set of replicate measurements and/or values expressed quantitatively by a statistic, such as SD or CV.

least squares regression – the method of statistically placing the location of the estimated line or curve among the data so that the sum of the squares of the distances of each data point from the line in the perpendicular direction from the x-axis (ie, parallel to the y-axis) is minimized; **NOTE:** It allows the direct algebraic computation of the coefficients and an estimate of their uncertainty.

measurand – quantity intended to be measured (JCGM 200:2012)[9]; **NOTE 1:** The specification of a measurand requires knowledge of the kind of quantity, description of the state of the phenomenon, body, or substance carrying the quantity, including any relevant component, and the chemical entities involved (JCGM 200:2012)[9]; **NOTE 2:** The measurement, including the measuring system and the conditions under which the measurement is carried out, might change the phenomenon, body, or substance such that the quantity being measured may differ from the measurand as defined. In this case, adequate correction is necessary (JCGM 200:2012)[9]; **EXAMPLE 1:** S-Creatinine concentration is frequently measured using an enzyme-based technique resulting in a color reaction. Results will be biased due to different specificity of the measurement procedures and different quantities measured; **EXAMPLE 2:** The length of a steel rod in equilibrium with the ambient temperature of 23°C will be different from the length at the specified temperature of 20°C, which is the measurand. In this case, a correction is necessary (JCGM 200:2012)[9]; **NOTE 3:** In chemistry, "analyte," or the name of a substance or compound, is a term sometimes used for "measurand." This usage is erroneous because these terms do not refer to quantities (JCGM 200:2012).[9]

measuring interval//working interval – set of values of quantities of the same kind that can be measured by a given measuring instrument or measuring system with specified instrumental measurement uncertainty, under defined conditions (JCGM 200:2012)[9]; **NOTE:** The lower limit of a measuring interval should not be confused with detection limit (JCGM 200:2012).[9]

ordinary linear regression (OLR) – least squares linear regression that usually refers to nonweighted least squares regression; **NOTE:** OLR may also be described as uniformly weighted ordinary least squares regression.

outlier – the observation in a sample, so far separated in value from the remainder as to suggest that it may be from a different population, or the result of an error in measurement; **NOTE 1:** The World Health Organization (WHO) defines this as "a number of a set of values that is inconsistent with the other numbers of the set" (WHO-BS/95.1793)[11]; **NOTE 2:** Statistical tests can be used to identify outliers, but the "common-sense" judgment using visual inspection of the data is often more effective.

Passing-Bablok regression – nonparametric procedures to estimate slope and intercept parameters from a measurement procedure comparison experiment.[12,13]

precision (measurement) – closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions (JCGM 200:2012)[9]; **NOTE 1:** Measurement precision is usually expressed numerically by measures of imprecision, such as SD, variance, or CV under the specified conditions of measurement (JCGM 200:2012)[9]; **NOTE 2:** The "specified conditions" can be, for example, repeatability conditions of measurement, intermediate precision conditions of measurement, or reproducibility conditions of measurement (see ISO 5725-3:1994)[14] (JCGM 200:2012)[9]; **NOTE 3:** Measurement precision is used to define measurement repeatability, intermediate measurement precision, and measurement reproducibility (JCGM 200:2012).[9]

replicate – a value resulting from a repeat analysis of the same specimen.

reproducibility (measurement) – measurement precision under reproducibility conditions of measurement (JCGM 200:2012).[9]

reproducibility condition (of measurement) – condition of measurement, out of a set of conditions that includes different locations, operators, measuring systems, and replicate measurements on the same or similar objects (JCGM 200:2012)[9]; **NOTE 1:** The different measuring systems may use different measurement procedures (JCGM 200:2012)[9]; **NOTE 2:** A specification should give the conditions changed and unchanged, to the extent practical (JCGM 200:2012).[9]

residual – the difference between a given data point and its predicted value. **NOTE:** As used in EP09, for evaluating a value predicted by a regression fit.

sample – one or more parts taken from a primary sample (ISO 15189)[15]; **NOTE 1:** For example, a volume of serum taken from a larger volume of serum (ISO 15189)[15]; **NOTE 2:** A sample is prepared from the patient specimen and used to obtain information by means of a specific laboratory test; **NOTE 3:** The system from which a sample is taken may not be of the same type as that of the measurand. For example, a given blood sample may serve for measurement of pH in plasma hemoglobin concentration in erythrocytes.

scatter plot//scatter diagram – a type of mathematical diagram using Cartesian coordinates to display values for two variables for a set of data; **NOTE 1:** The data are displayed as a collection of points, each having the value of one variable determining the position on the horizontal axis and the value of the other variable determining the position on the vertical axis; **NOTE 2:** Also called scatter chart, scattergram, scatter diagram, or scatter graph.

trueness (measurement) – closeness of agreement between the average of an infinite number of replicate measured quantity values and a reference quantity value (JCGM 200:2012)[9]; **NOTE 1:** The measure of trueness is usually expressed in terms of bias (ISO 5725-1)[16]; **NOTE 2:** Measurement trueness is inversely related to systematic measurement error, but is not related to random measurement error (JCGM 200:2012)[9]; **NOTE 3:** Measurement accuracy should not be used for "measurement trueness" and vice versa (JCGM 200:2012).[9]

validation – verification, where the specified requirements are adequate for an intended use (JCGM 200:2012)[9]; **EXAMPLE:** A measurement procedure, ordinarily used for the measurement of mass concentration of nitrogen in water, may be validated also for measurement in human serum (JCGM 200:2012).[9] **NOTE:** An external claims verification conducted by the manufacturer for the purpose of validating that the measurement procedure is fit for the purpose of quantitatively determining the concentration of a measurand in a clinical laboratory setting is an example of a validation.

verification – provision of objective evidence that a given item fulfills specified requirements (JCGM 200:2012)[9]; **EXAMPLE 1:** Confirmation that a given reference material as claimed is homogeneous for

the quantity value and measurement procedure concerned, down to a measurement portion having a mass of 10 mg (JCGM 200:2012)[9]; **EXAMPLE 2:** Confirmation that performance properties or legal requirements of a measuring system are achieved (JCGM 200:2012)[9]; **EXAMPLE 3:** Confirmation that a target measurement uncertainty can be met (JCGM 200:2012)[9]; **NOTE 1:** When applicable, measurement uncertainty should be taken into consideration (JCGM 200:2012)[9]; **NOTE 2:** The item may be, eg, a process, measurement procedure, material, compound, or measuring system (JCGM 200:2012)[9]; **NOTE 3:** The specified requirements may be, eg, that a manufacturer's specifications are met (JCGM 200:2012)[9]; **NOTE 4:** Verification should not be confused with calibration. Not every verification is a validation (JCGM 200:2012).[9]

weighted regression – a parametric regression analysis technique that weights the influence of individual patient results based upon a predefined criterion; **NOTE 1:** Weighted regression is often applied to both least squares (*y* results weighted) and Deming (*x* and *y* results weighted) regressions; **NOTE 2:** The weighting scheme typically assigns influence to a result that is inversely related to its expected variance.

4.3 Symbols Used in the Text

The following symbols are used in this document. See Table 1 for a list of terms used to describe the two measurement procedures being compared.

X	*X* characteristic of a comparative measurement procedure
Y	*Y* characteristic of a candidate measurement procedure
N	total number of samples
r	correlation coefficient
x	observation from comparative measurement procedure
x_i	estimate of comparative measurement procedure's value for sample number *i*. This may either be a single comparative measurement procedure replicate from a sample or the average of multiple comparative measurement procedure replicates from that sample.
$\bar{x}$	average of the *x* observations
y	observation from candidate measurement procedure
y_i	estimate of candidate measurement procedure's value for sample number *i*. This may either be a single candidate measurement procedure replicate from a sample or the average of multiple candidate measurement procedure replicates from that sample.
$\bar{y}$	average of the y observations
d_i	difference between comparative and candidate measurement procedures for sample number *i*
$\bar{d}$	average of sample result differences between comparative and candidate measurement procedures
z_i	position on horizontal axis of a difference plot for sample number *i*
b	slope of the regression line
a	*y* intercept of the regression line
$\hat{Y}$	predicted value for candidate measurement procedure
S_{yx}	standard deviation of residuals of regression (standard error of estimate)
X_c	selected concentration of comparative measurement procedure (eg, medical decision level)
$\hat{B}_c$	estimate of predicted bias at concentration X_c
B_c	true bias at concentration X_c

4.4 Abbreviations and Acronyms

CI	confidence interval
CV	coefficient of variation

ESD	extreme studentized deviate
IVD	*in vitro* diagnostic
OLR	ordinary linear regression
QC	quality control
SD	standard deviation
SE	standard error
VIM	*Vocabulaire international de métrologie; International vocabulary of metrology – Basic and general concepts and associated terms* (JCGM 200:2012)
WLS	weighted least squares

5 Measurement Procedure–Familiarization Period

The operators of both the candidate and the comparative measurement procedures must be familiar with the following:

- Operation of the instrument systems and procedural steps to perform the measurement procedures
- Maintenance procedures of these instrument systems
- Methods of sample preparation
- Calibration and measurement procedure quality monitoring functions

Manufacturers' training programs, when offered, can be a part of the familiarization period. Clinical laboratory personnel must set up and operate the required instrument systems long enough to ensure that the operators understand all procedures and can properly operate them. Five days are recommended for the measurement procedure–familiarization period. For extremely simple instrument systems, a shorter period can suffice; for complex, multichannel instrument systems, a longer period can be required.

The operators should practice analyzing unmodified patient sample materials to bring to their attention all possible contingencies (eg, error flags, error correction) that might arise during routine operation of either instrument system. Data collected during this period can be used as objective evidence that the personnel are qualified to perform the measurement procedures being compared, but not as study data. The measurement procedure–familiarization period is not complete until the operators can perform the measurement procedures with confidence.

The familiarization period is optional for a manufacturer who is conducting the study under its own standard procedures as long as the personnel conducting the study have been trained on these procedures and have experience running both measurement procedures.

6 Measurement Procedure Comparison Studies

This section covers experimental considerations for the manufacturer. In most cases, these considerations are in common with those for the clinical laboratory. The considerations unique to the clinical laboratory are covered in Section 7.

When a manufacturer or research laboratory creates an IVD measurement procedure, a comparison study is typically performed to help determine whether this candidate measurement procedure successfully quantifies the measurand. Successful quantification is determined when the average bias over the measuring interval and possibly the bias measured at specified measurand concentrations are within preestablished specifications when comparing a candidate to a comparative measurement procedure. Ideally, the comparative measurement procedure for such a comparison is an accepted standard or reference measurement procedure with low imprecision and bias. However, the comparative measurement procedure is often a commercially available measurement procedure from another manufacturer, which may have limitations with regard to imprecision, measuring interval, linearity, and specificity for the

measurand. This may be especially true for an automated candidate measurement procedure that is being designed to replace a manual procedure. In this case, the comparative measurement procedure may have significantly higher imprecision than the candidate.

Regardless of the imprecision inherent in the comparative measurement procedure, during research and development process, the manufacturer may perform multiple measurement procedure comparisons between it and the candidate measurement procedure to create a robust calibration scheme. Such testing is beyond the scope of this document.

6.1 Study Samples

The goal for any measurement comparison study is to determine the relationship between the candidate and comparative procedures using unmodified samples that cover their entire measuring intervals.

Collect and handle patient samples according to accepted laboratory practice and manufacturers' recommendations. Any clinical, demographic, or analytical (eg, hemolysis, icterus, and lipemia/turbidity indices) information available on the patients providing the samples should be retained.

If it is desired to perform replicate determinations or to retain enough sample for possible retesting, and the required volume of a sample cannot be obtained from a single patient, then make "minipools" by mixing samples from multiple patients (when possible, use two) with approximately the same concentration of measurand and similar disease histories whenever possible. In cases in which specific measurand concentrations are desirable but not available for spanning the entire measuring interval, two samples, with disparate concentrations of measurand can be pooled. If the samples are whole blood, mixing requires serological compatibility. Any such modified patient samples must be noted in any listing of results, and ideally also in any graphical representation of the data. Ideally, modified samples comprise a small portion of the samples in the study (eg, no more than 20%).

NOTE: The process of pooling can mask sample-to-sample characteristics by averaging out unique or sample-specific influences and thus can lead to an optimistic picture of the comparability of the two measurement procedures. To minimize this effect, material from an individual patient sample should be used in no more than one minipool sample aliquot to be tested. However, in cases of rare diseases or specimens, it may be necessary to use an individual patient sample to prepare multiple aliquots at different measurand concentrations.

When the biological specimens under study are tissue samples rather than body fluid aliquots, heterogeneity is a concern. Handling of such samples is covered in Appendix A1 of CLSI document I/LA28.[17] Their characteristics are not amenable to the measurement of multiple replicates and certainly prohibit the pooling of multiple samples.

Occasionally, upon sample collection there are many more low concentration samples than ones of higher concentration. In such cases, the first option is to see if the sample set can be supplemented with additional high concentration samples. For some measurement procedures, the incidence rate of such high samples is so low that such supplementation is not feasible. A visualization technique for such datasets is provided in Section 8.3.4.

When unmodified patient samples are difficult to obtain at specific portions of the measuring interval, other options can be investigated such as dilution or depletion to obtain lower concentrations, or spiking analyte into unmodified samples to obtain high concentration samples. However, such techniques should only be used as a last resort because modified samples are likely to have commutability limitations between two different measurement procedures. In fact, modified samples have been shown to have commutability limitations for different reagent lots for the same measurement procedure.[18] If such

techniques are used, they should be labeled as such in any plots where they are presented and analysis should be performed both with and without such samples.

In any study comparing candidate to comparative measurement procedures, only one of the sample types (eg, serum or plasma) recommended for both measurement procedures should be used. The sample type used in the comparison should be stated.

6.1.1 Storage

The accumulation of samples for the measurement procedure comparison study may require storage of these samples, especially if the two measurement procedures are in different locations. The clinical laboratory should ensure that variation related to damage or deterioration due to transport or storage does not impact either measurement procedure. To the extent possible, the split samples used on the candidate and comparative measurement procedures should be treated in a similar manner with respect to storage and handling.

6.2 Comparative Measurement Procedure

This experiment gives an estimate of the bias between two measurement procedures and estimates for bias, at any specified concentration. The estimates of the concentration for a sample from the comparative measurement procedure should ideally have the following characteristics:

- Have lower uncertainty than estimates of the concentration for that same sample from the candidate measurement procedure, which can often be achieved by averaging over replicates, if needed.

- Be free from known interferences, whenever possible. This should be true for both the comparative and the candidate measurement procedures.

- Use the same units as the candidate measurement procedure or have the ability to be converted to the same units.

- Be traceable to standards or reference measurement procedures, whenever possible.

This experiment does not segregate the various sources of bias into those coming from each of the measurement procedures being compared. (See CLSI document EP14[19] for information on detection of matrix interference.) Interference effects may contribute as much as imprecision effects to a difference between measurement procedures in any given sample. (Proper characterization of interference effects on each measurement procedure can be determined by a separate experiment; see CLSI document EP07.[5])

6.2.1 Interval of Study Measurements

The goal for manufacturers of both establishment and claims verification (validation) studies is to determine the relationship between the candidate and comparative measurement procedures over the broadest interval possible. Every attempt must be made to collect samples that cover the entire measuring interval of the measurement procedure rather than just the reference interval or the clinical decision points.

However, the measurement of measurand concentration is restricted by the analytical measuring intervals of the two measurement procedures (ie, where they overlap). The analytical measuring interval is the measurand concentration interval claimed by the manufacturer to provide acceptable performance. Ideally, the measuring interval of the comparative measurement procedure will be at least as wide as the measuring interval of the candidate measurement procedure so that bias at the limits of the analytical measuring interval can be compared.

Sometimes a candidate measurement procedure is developed in order to meet an unmet clinical need for an extended interval of measurand values. If a reference measurement procedure is available that covers this extended interval, it can be used as the comparison. Often, however, there is no reference method and the comparative measurement procedure has a restricted measuring interval. In such a case, if the dilution capabilities of the comparative measurement procedure have been verified, then unmodified high concentration samples on the candidate measurement procedure can be compared to diluted samples on the comparative measurement procedure to cover the extended measuring interval.

6.3 Number of Samples

It is recommended that, for establishment and claims verification studies, manufacturers use at least 100 samples that meet the criteria stated in Sections 6.1 and 6.2.

If two clinically relevant populations have been shown to provide different relationships between the candidate and comparative measurement procedures, then each such population will require a study of the recommended 100 samples. Such differences may be stated in the manufacturer's product labeling or in the clinical literature. As an example, two immunoassays may exhibit one relationship on samples from pregnant female patients and another relationship on samples from male patients, due perhaps to different concentrations of cross-reacting substance in the two populations that have different effects on the two measurement procedures. In another example, two measurement procedures for parathyroid hormone may diverge markedly from one another on samples from dialysis patients even when they agree closely on samples from patients with normal kidney function.

6.3.1 Measurement Replicates

Obtain a sufficient amount of each sample so that the number of replicates specified for the candidate and comparative measurement procedures can be run.

For a manufacturer's establishment of bias performance, the matched sample-to-sample bias comparison requires that the average of each sample concentration be used. Therefore, if multiple replicates are available, they should be averaged to estimate each sample concentration. The underlying assumption behind this averaging of results is that both replicates, from each measurement procedure, are attempting to measure the same, unchanging quantity for that sample[20] and that an average therefore reduces the uncertainty (standard error) of the estimate for that sample. If three or more replicates are available, the use of median rather than average for the sample value estimate is a reasonable alternative.

When the manufacturer verifies that its requirements are being met through a claims verification study (validation), then the measurement of a sample's concentration must be derived as it would be during the intended use of the measurement procedure. Therefore, if a candidate measurement procedure uses one replicate to produce patient results, then a typical strategy is to collect one replicate per sample for this procedure in the manufacturer's validation (see Section 2.2.2). If, however, a measurement procedure (ie, a manual measurement procedure) requires that two replicates be averaged to get a patient result (eg, enzyme-linked immunosorbent measurement procedure), then the average of two replicates should be used in such a validation. For such a study, the manufacturer may or may not have more leeway to collect and average multiple replicates from the comparative measurement procedure. The use of multiple replicates for either procedure is reasonable only in cases in which sample results are demonstrated not to be dependent on replicate order.

6.4 Factors Affecting the Measurement Procedure Comparison

Many experimental factors affect the bias estimate from a measurement procedure comparison study. Random factors such as within-run and between-run (including between-day) variability play a role. It is generally expected that such random factors will not create a systematic shift in bias but only affect the

variability of the bias estimate. Therefore, an experimental design that increases the replication over such factors (eg, number of runs, number of days, or number of replicates per run) will decrease the bias estimate uncertainty and thus the confidence interval (CI) for the estimate of bias.

Other factors can cause systematic shifts of the bias estimate. These may include shifts due to operator, calibrations, instruments, reagent lots, and calibrator lots. The manufacturer should have an idea of which factors provide the highest potential for such bias shifts. The ideal IVD measurement procedure uses the calibration procedure to eliminate such shifts, but not all IVD measurement procedures are ideal. If such a factor is significant, its effect can also be moderated by increasing the number of instances of that factor in the study (eg, increasing the number of calibrations).

6.4.1 Establishment of Relationship Study Design

The goal of a measurement procedure comparison study to establish a relationship is to determine the candidate measurement procedure's bias relative to the comparative measurement procedure. In some cases, comparisons are to a reference or other comparative measurement procedure with an expected bias to the candidate measurement procedure. The manufacturer will have determined customer needs and performed testing on the candidate measurement procedure during its development process. Both of these sources of information should be used to determine an expectation for bias, which may be different than zero, and the factors that should be considered for such a study.

An establishment study is an analytical study typically performed at the manufacturer's site that answers the question, "What is the relationship between the candidate and the comparative measurement procedures?" It is up to the manufacturer to determine the rigor of the study based on knowledge accumulated on the candidate measurement procedure being tested. During assay development, the manufacturer will have determined the potential influence of various factors on candidate measurement procedure bias. Because the manufacturer will later verify the claim established by this study, the manufacturer can determine for itself how the establishment study is designed.

A study, for example, could be conducted over three to five days with a relatively equivalent number and concentration distribution of samples run each day. Besides the day factor, the manufacturer may choose to include additional factors in the study design such as reagent lot, calibrator lot, calibration, instrument, and operator.

If the comparative measurement procedure is a well-controlled standard or reference measurement procedure, then none of these factors need to be considered for this measurement procedure, because each of the individual comparative results would be defined as true (within expected random error). Often the manufacturer may be hampered by the lack of information on the variability for the comparative measurement procedure and the lack of access to multiple instruments and reagent lots. In either case, each subset of comparative measurement procedure results will ideally be generated proximal to the same time as possible to the same subset of results generated by the candidate measurement procedure.

Averaging over multiple replicates can reduce the imprecision of an estimate of bias for an establishment study. However, the sample volume available will frequently limit the number of replicates. For such a study, if three or more replicates are available, the use of median for the sample value estimate rather than average is a reasonable alternative.

6.4.2 Manufacturer's Claims Verification Study (Validation) Design

The goal of a manufacturer's claims verification study, usually conducted at one or more clinical sites, is to show that the candidate measurement procedure can meet bias specifications while run under typical clinical laboratory conditions. A measurement procedure can also be shown to be acceptable if a known difference (non-zero) is confirmed. This can be seen as a validation that the measurement procedure is fit

for the purpose of quantitatively determining the concentration of a measurand in a clinical laboratory setting.

The study should be performed over a minimum of three to five days. Other factors similar to those mentioned for an establishment of relationship study can be considered in conducting such a validation experiment. Typically, however, complex study designs will not be feasible in an experiment conducted at an external clinical laboratory. However, performing multiple calibrations during the study and combining data from more than one clinical laboratory may increase the robustness of bias estimates.

Both measurement procedures should be run as described in their instructions for use. Therefore, if only one replicate is used for providing a measurement, then a typical strategy is to collect only one replicate to protect against the possibility that subsequent replicates are not exchangeable with the first.[21]

Using a single replicate per instructions for use also ensures the study validates that a candidate measurement procedure is fit for purpose. Statistically, this fact does not invalidate the estimation of bias using averages of multiple replicates. Provided each replicate represents equivalent information (ie, order has no influence), each represents a result from the procedure per the instructions for use, so averaging of the results on multiple replicates simply improves bias estimation.

If multiple replicates are being collected for either procedure, then, as mentioned in Section 6.4.1, either the average or median results may be used in the calculation of the sample value for that measurement procedure. Generally, the average is the summary of choice, unless use of the median is needed to offset the effect of skewness in the distribution of replicate values.

The resulting estimates of bias (average bias over the measuring interval or bias at specified concentrations) should be compared to acceptance criteria established before study initiation.

6.5 Sample Sequence

During a prospective study, sample sets may be run as they are collected or they may be organized into sets (eg, that cover the measuring interval) for later daily batch testing. Within each set, all samples should be run in random sequence for both measurement procedures. If, within each set of samples, individual replicates are run independently (eg, via a manual method), then randomize the replicates, as well. For a random access, automated instrumentation, there is no need to randomize individual replicates within each set of samples.

6.6 Time and Duration

For a given sample, measurement by the comparative and candidate procedures should occur within a time span consistent with measurand stability. If possible, use samples drawn the day of the analysis. If stored samples are used, make sure they were all stored in a manner that ensures their stability and meets the stated requirements of both the candidate and the comparative measurement procedures. Store samples in the same manner for both procedures to avoid introducing storage condition as a variable.

6.7 Inspection of Data During Collection

Inspect data during its collection, because blunders of data collection such as misalignment of sample names with results from a sample are easier to catch during testing than after testing is completed. At this stage, such errors with assignable cause can be corrected with no controversy. Later, it may be difficult to determine the cause of discrepant data, and therefore difficult to correct such an error. Multiple replicates of each measurement make the determination of cause much easier, whether at this stage or later upon data visualization.

All replicates should be retained and presented in a table, along with the averages, if the average is to be used as that sample's concentration estimate. If data are manually transcribed in the table, it is recommended to inspect each entry against the original instrument value to detect possible transcription errors. This table should be inspected for obvious outliers. Such outliers should be investigated for potential errors caused by instrumentation, human, or procedural errors. This is an initial review that may not detect all such errors. Later, when the data are inspected visually (see Section 8), this table can again be reviewed to see if individual replicate errors are causing errors in sample concentration estimates.

Document data collected during a time when an instrument system indicates that an error condition exists, but do not include them in the final data analysis.

Record any data for which the operator can document that an error was made, but do not include them in the final analysis.

6.8 Quality Control

Follow appropriate QC procedures during the experiment. Keep control charts, and repeat any run that appears out of control on either measurement procedure until the required number of samples is obtained.

6.9 Documentation of Rejected Data

Carefully document and retain a record of any situation that requires the rejection of data (replicates or samples) along with any discovered causes and problems.

7 Considerations for Clinical Laboratories

The purpose of a measurement procedure comparison conducted by a clinical laboratory can either be the verification of bias performance claimed by the manufacturer or independent quantification of bias. Either goal can be achieved through the use of this guideline.

The considerations covered in Section 6 for manufacturers also cover studies conducted by clinical laboratories, except for the unique considerations covered below in this section.

7.1 Comparative Measurement Procedure

For the comparative measurement procedure, use the clinical laboratory's current measurement procedure, the measurement procedure used by the manufacturer in the labeled claims, or a recognized reference measurement procedure. The former is the most likely because the clinical laboratory will usually want to understand the difference between a measurement procedure to be introduced and the one they have been using. The laboratory should be aware that any comparative method, other than a reference measurement procedure, may have some vulnerability to interfering substances and matrix effects.

7.2 Number of Samples

Clinical laboratories should analyze at least 40 samples that meet the criteria stated in Sections 6.1 and 6.2 to establish the bias between measurement procedures. More samples will improve the confidence in the statistical estimates and increase the opportunity to incorporate the effects of unexpected interfering substances (individual idiosyncratic biases).

7.2.1 Measurement Replicates

For the clinical laboratory, single measurements per procedure are acceptable if deemed appropriate by the laboratory director.

For measurement procedure introduction into a clinical laboratory (see Section 2.2.3), the matched sample-to-sample bias comparison requires that the best estimate of each sample concentration be used. Therefore, if multiple replicates are available, they should be averaged (or the median taken) to estimate each sample concentration. The underlying assumption behind this averaging of results is that the replicates, from each measurement procedure, are attempting to measure the same, unchanging quantity for that sample[20] and that an average therefore reduces the uncertainty of that estimate of sample concentration.

7.3 Calibration and Procedure Control

Calibrating both the candidate and comparative measurement procedures at the start of the study is recommended to ensure that each is in conformance with all QC parameters. If necessary, recalibrate as indicated in the instructions for use or the laboratory's operating procedure for either measurement procedure.

8 Visual Data Review

Having completed the collection of data for analysis, the next step is the visual review of the data. Such a review is useful to evaluate whether the desired interval is adequately covered, to screen for the presence of exceptional values or outliers, to get an initial understanding of the difference between the measurement procedures, and to determine how best to characterize the variability of these differences across the overlapping interval of measurements provided by the two measurement procedures. Two robust and flexible tools for this review are the scatter plot and the difference plot.[22]

The examples in this section and Section 9 are pedagogical for the purpose of describing methods for visual data review and quantitative analysis. They are not intended to imply acceptance of particular aspects of study design. Study design is discussed elsewhere (see Sections 5, 6, and 7) for the three types of studies considered in the document. In particular, the examples do not have the number of samples required for manufacturer studies, and in many cases the number of samples at higher concentrations may be smaller than would be found in an optimal design. Such examples are offered because many of the difficulties that various visualization and analysis techniques try to mitigate are also mitigated by a larger, more evenly spaced sample set. The data tables for most of the examples are provided in Appendixes I and J. They will be denoted as Table I1 or J1, J2, etc., if the reader wishes to review them.

8.1 Scatter Plots

A scatter plot presents the results of a measurement procedure comparison, with the comparative measurement procedure on the x-axis and the candidate measurement procedure on the y-axis. These plots can be visually inspected to determine the underlying variability characteristics of this relationship.

Two measurement procedure comparisons are presented below. Figure 1 displays a consistent amount of variability (SD) across the measuring interval (from Table J1). Figure 2 displays a variability that is proportional to concentration, which can be expressed as a constant CV, where CV = SD / concentration (from Table J2; see Appendix J for these and other tables that contain data for the figures that follow). This is seen as a cone of data converging to a point at the lower left-hand corner and opening up to the upper right.

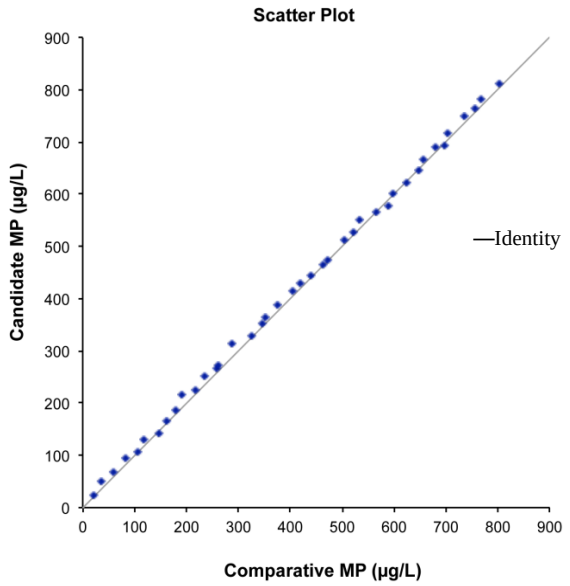

Figure 1. Scatter Plot With Constant SD (From Table J1)
Abbreviation: MP, measurement procedure.

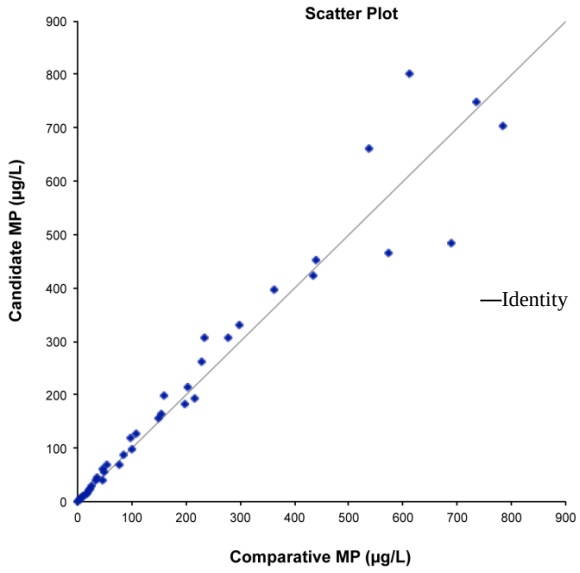

Figure 2. Scatter Plot With Constant CV (From Table J2)
Abbreviation: MP, measurement procedure.

8.2 Difference Plots

A difference plot[23] presents the results of a measurement procedure comparison, with the measurand concentration on the horizontal axis and the difference between the candidate and the comparative measurement procedures on the vertical axis (see Figure 3 below). Bland-Altman[22] is an example of a difference plot. Such plots can be visually inspected to determine the underlying variability characteristics of this relationship.

The user must select from four types of difference plots based on two factors. The first factor is determined by whether the user wishes to see the comparative method as the truth against which the candidate method is compared or to see the average of the two methods as the best estimate of the true value for a sample. In the first case, the horizontal axis of the plot is the result from the comparative measurement procedure.[24] In the second case, advocated by Bland and Altman,[25] the horizontal axis is the average of the two measurement procedures' results.

When a reference measurement procedure is the comparative measurement procedure, its results should be used on the horizontal axis. A manufacturer may wish to use the most common measurement procedure as the comparative measurement procedure. In this case, when the comparative measurement procedure is not considered a reference, the average result of the two measurement procedures (candidate and comparative) may be used on the horizontal axis for data visualization.

A clinical laboratory may wish to use its current measurement procedure as the comparative measurement procedure and may consider it to be a reference because the goal is to compare the known behavior of its current procedure against the unknown candidate measurement procedure. In this case, the results for the comparative measurement procedure should be used on the horizontal axis.

The second factor is whether the variability of the differences between the two measurement procedures is constant or proportional to the concentration on the horizontal axis. In the first instance, the magnitude of the difference is assumed to be essentially the same across the entire interval of concentrations (see Figure 1). In the second instance, the magnitude of the difference is assumed to be proportional to

concentration (see Figure 2). Because this characteristic of the relationship may not be known beforehand, it is suggested that both reporting unit and percent difference plots be created and inspected (eg, Figures 3A and 3B). The equations required to create each of the four plots are provided in Table 3.

Table 3. Formulas for Creating Difference Plots

Horizontal Axis (*z*)	Vertical Axis	
	Difference (*d*) Is Constant (Constant SD)	**Difference (*d*) Is Proportional to Concentration (Constant CV)**
Comparative measurement procedure	$z_i = \text{concentration} = x_i$ $d_i = \text{difference} = y_i - x_i$ (1)	$z_i = x_i$ $d_i = (y_i - x_i)/x_i$ (2)
Average of the two procedures	$z_i = (x_i + y_i)/2$ $d_i = y_i - x_i$ (3)	$z_i = (x_i + y_i)/2$ $d_i = (y_i - x_i)/[(x_i + y_i)/2]$ (4)

Abbreviations: CV, coefficient of variation; SD, standard deviation.

Parameters: x_i is the result of the comparative measurement procedure for patient sample number *i;*
y_i is the result of the candidate measurement procedure for patient sample number *i;*
(z_i, d_i) are the resultant coordinates on the difference plot for patient sample number *i.*

8.3 Inspect Plots for Underlying Characteristics

The optimal technique used to determine the bias between the candidate and comparative measurement procedures is highly dependent upon whether the data meet specific underlying assumptions. First determine whether the variability of differences between the two measurement procedures is constant or proportional to concentration.

8.3.1 Constant Difference Variability (Constant Standard Deviation)

If the variability of the differences between the candidate and comparative measurement procedures is constant, the two plots will appear as they do in Figure 3:

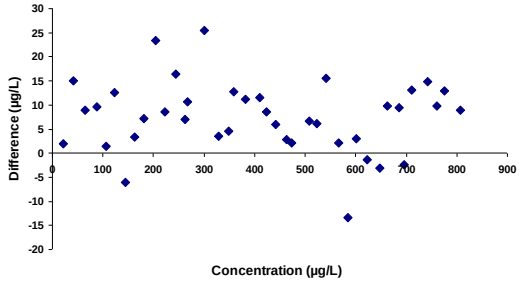

Figure 3A. Reporting Units Difference Plot

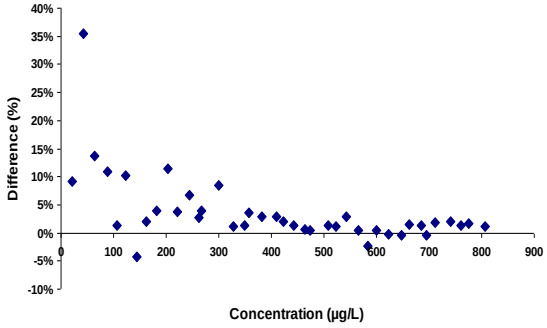

Figure 3B. Percent Difference Plot

Figure 3. Constant Difference Variability Between Measurement Procedures (From Table J1)

Note that the spread of the differences remains consistent across the range of concentration on the reporting units difference plot, but that the spread of the differences increases significantly with decreasing concentration on the percent difference plot. When the difference is constant across the interval of concentration, the reporting unit difference plot provides the better representation of the difference between measurement procedures.

8.3.2 Proportional Difference Variability (Constant Coefficient of Variation)

If the variability of the differences between the candidate and comparative measurement procedures is proportional to concentration, the appearance of the two plots will appear as it does in Figure 4:

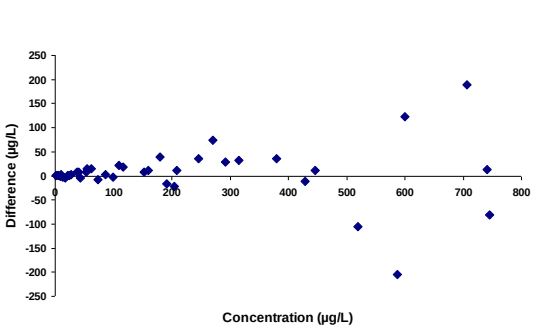

Figure 4A. Reporting Units Difference Plot

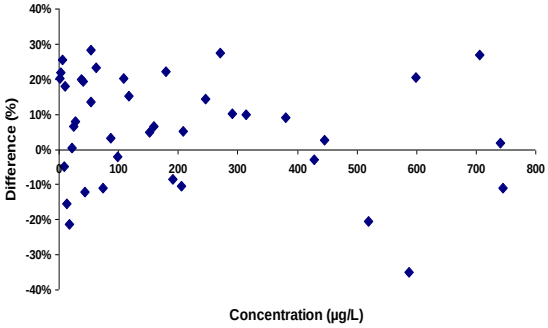

Figure 4B. Percent Difference Plot

Figure 4. Proportional Difference Variability Between Measurement Procedures (From Table J2)

In this case, the situation is reversed from Figure 3. The reporting units difference plot provides a spread of differences that gets wider as the horizontal axis concentrations get higher, while the percent difference plot provides a spread that is consistent over the horizontal axis. In this case, the percent difference plot provides the better representation of the difference between measurement procedures.

8.3.3 Mixed Difference Variability (Standard Deviation and Coefficient of Variation)

Often a measurement procedure will exhibit a mixture of these two characteristics with the differences being constant at low concentrations and proportional to concentration at higher concentrations. Such a measurement procedure is shown in Figure 5.

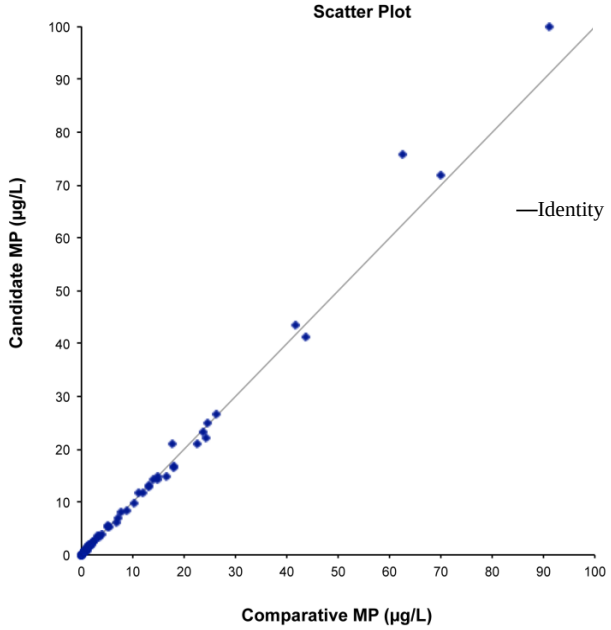

Figure 5A. Scatter Plot

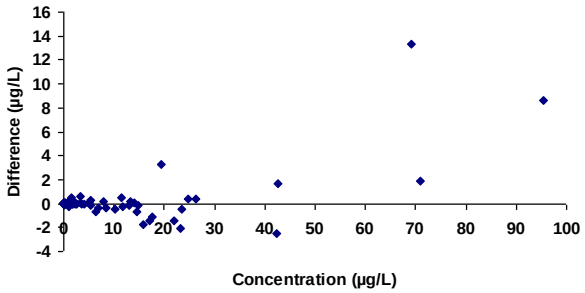

Figure 5B. Reporting Units Difference Plot

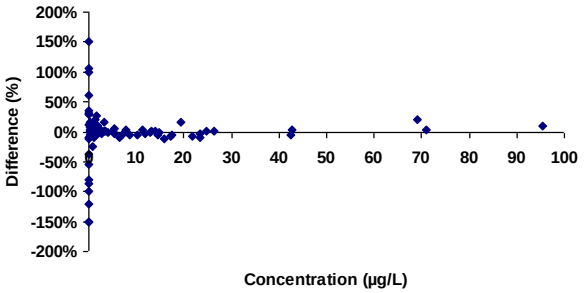

Figure 5C. Percent Difference Plot

Figure 5. Mixed Difference Variability Between Measurement Procedures (From Table I1)
Abbreviation: MP, measurement procedure.

Through inspection of the plot (or the underlying data used to generate the plot), an estimate may be obtained of the concentration at which the relationship changes from a constant difference to a proportional difference. The concentration at which the relationship changes from constant to proportional variability can be estimated by formal statistical analysis (called change point analysis), but is beyond the scope of this guideline.

The samples in this example and in Figure 6A below do not evenly cover the measuring interval. More samples at higher concentrations should be collected to meet the needs of a manufacturer's study.

8.3.4 Ranked Order Difference Plot

Regardless of how diligently samples are collected, occasionally a final dataset will have subintervals where there are few data points. Most commonly these will be at higher concentrations where only patients with relatively rare disease states will provide results. In such cases, usually in measurement procedures with proportional variability, the individual points will be widely separated at higher concentrations (see Figure 6), making determination of variability characteristics more challenging.

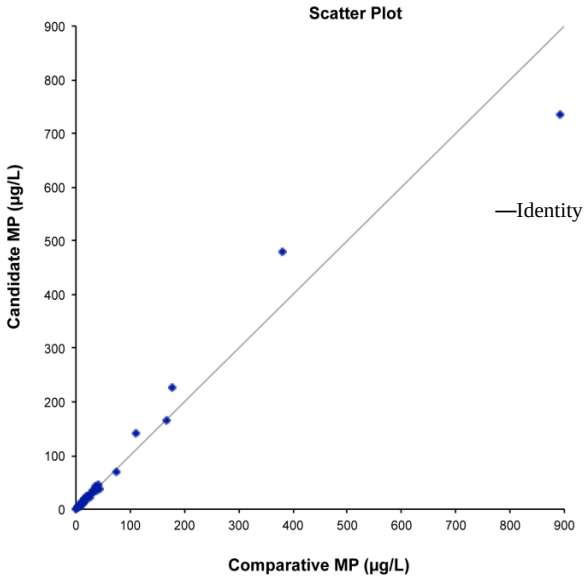

Figure 6A. Scatter Plot

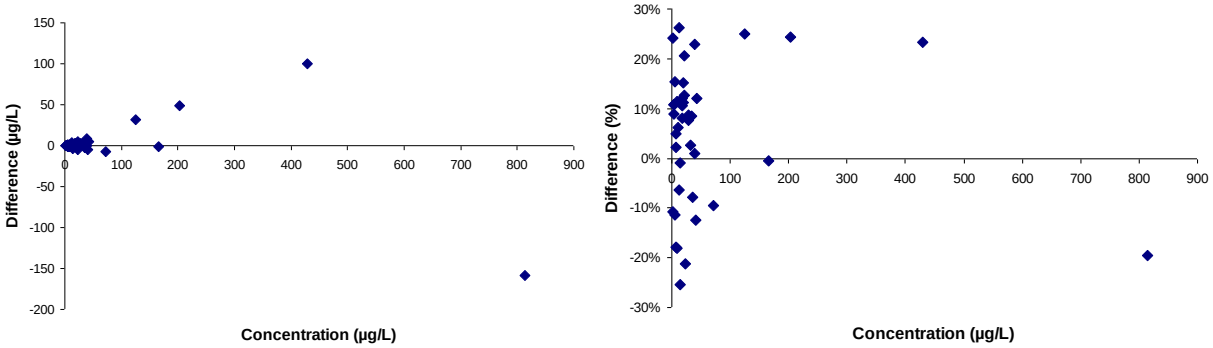

Figure 6B. Reporting Units Difference Plot **Figure 6C. Percent Difference Plot**

Figure 6. Proportional Difference Variability Between Measurement Procedures (From Table J3)
Abbreviation: MP, measurement procedure.

A ranked order difference plot can help in visualizing such a dataset. The first step in creating such a plot is to rank the N data points by concentration from lowest to highest and assign them numbers from 1 to N in that order. The options for the vertical axis are the same as those shown in Table 3.

The formulas for creating ranked order difference plots are provided in Table 4.

Table 4. Formulas for Creating Ranked Order Difference Plots

Horizontal Axis (z)	Vertical Axis: Difference (d) Is Constant (Constant SD)	Vertical Axis: Difference (d) Is Proportional to Concentration (Constant CV)
Samples ranked by comparative measurement procedure results	z_k = Rank(x_i) $d_k = y_k - x_k$ (5)	z_k = Rank(x_i) $d_k = (y_k - x_k)/x_k$ (6)
Samples ranked by average of the two procedures	z_k = Rank($[x_i - y_i)]/2$) $d_k = y_k - x_k$ (7)	z_k = Rank($[x_i - y_i)]/2$) $d_k = (y_k - x_k)/[(x_k + y_k)/2]$ (8)

Abbreviations: CV, coefficient of variation; SD, standard deviation.

The parameter k is the rank order of the samples (ranked by concentration and then by order of data collection in the case of ties in concentration)

The proportional difference plot (see Figure 6C) is shown below in Figure 6D, with this optional horizontal axis.

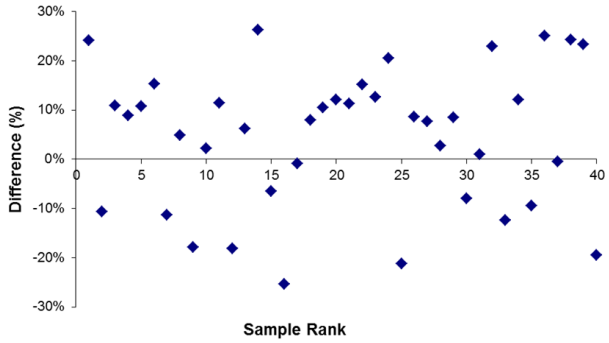

Figure 6D. Difference Plot With Ranked Order Horizontal Axis (From Table J3)

The constant proportional difference is quite apparent when viewed in this way. Such a technique is also useful for understanding where the relationship changes from a constant difference to a proportional difference. The mixed difference example from Figure 5 is presented in Figure 7, using the ranked sample number technique.

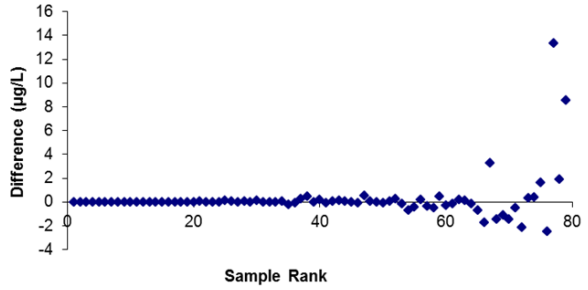

Figure 7A. Reporting Units Difference Plot

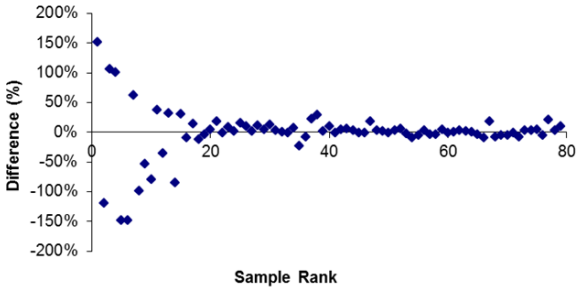

Figure 7B. Percent Difference Plot

Figure 7. Mixed Difference Variability Between Measurement Procedures (From Table I1)

Note that the point at which the mixed difference plot changes from constant to proportional is much easier to determine with this optional horizontal axis. The goal of picking horizontal (displaying the *z*-values) and vertical (displaying the *d*-values) axes is to create a difference plot that will most readily display the relationship between the two measurement procedures so that the underlying assumptions can be visually inspected.

Inspecting Figure 7A, there appears to be a constant variability from sample 1 through 35 to perhaps 55. Figure 7B displays relatively consistent proportional differences from sample 79 down to sample 40 to perhaps 30. It is important to note that portrayal of data in a ranked order plot as shown in Figure 7 enables one to better determine the region where this change occurs than does portrayal in a more typical difference plot, as shown in Figure 5.

8.3.5 When Bias Changes With Concentration

In some cases, bias may vary across the measuring interval irrespective of the variability pattern. In Figure 8, the variability of the differences is consistent across the measuring interval, but the magnitude of the difference (bias) changes in a linear fashion.

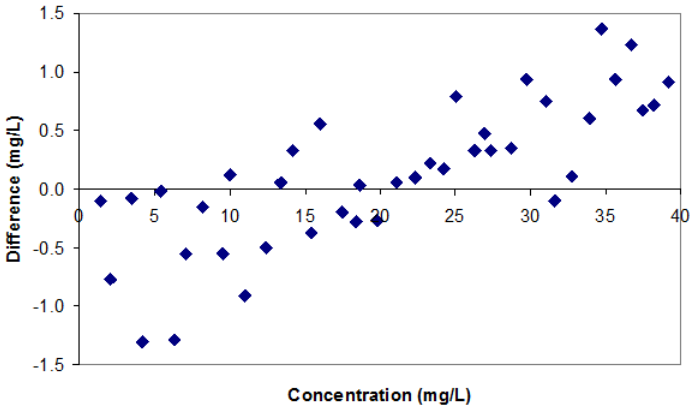

Figure 8. Bias Changes With Concentration

8.3.6 Nonlinear Relationship

Figure 9 illustrates a special case of nonconstant differences. In this dataset the variabilities of the differences are proportional to concentration; however, the magnitudes of the difference between the candidate and comparative measurement procedures change across the measuring interval in a nonlinear manner.

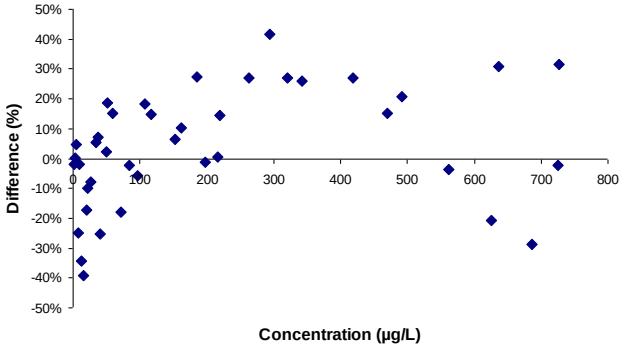

Figure 9. Nonlinear Relationship

8.3.7 Visualizing Anomalous Results

Both scatter plots and difference plots as shown in Figure 10 are useful in visualizing anomalous results.

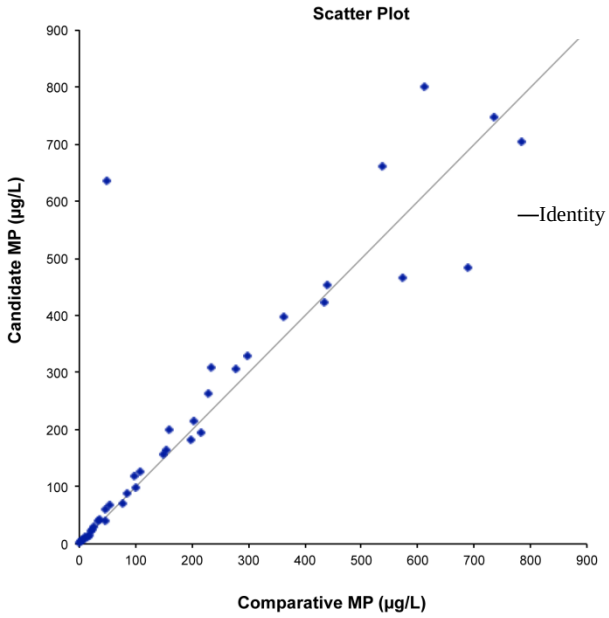

Figure 10A. Scatter Plot

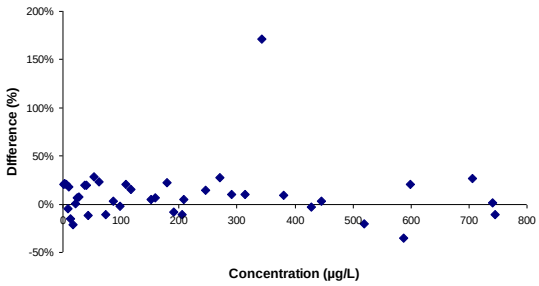

Figure 10B. Percent Difference Plot: Horizontal Axis = Average of Both Procedures

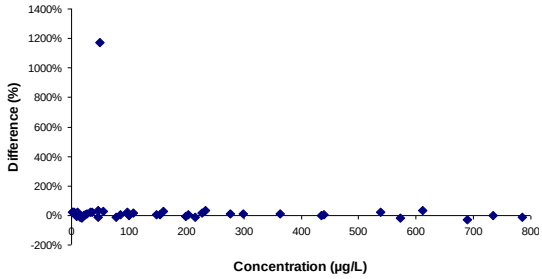

Figure 10C. Percent Difference Plot: Horizontal Axis = Comparative Procedure

Figure 10. Single, Outlying Point (From Table J4)
Abbreviation: MP, measurement procedure.

The performance of difference plots to demonstrate anomalous data using the two horizontal axis options is shown in Figure 10. Figure 10B shows each point using its distance from the identity line in the direction perpendicular to that line. This is similar to the way an outlying point is typically perceived on a scatter plot (Figure 10A). Figure 10C also shows the distance from the identity line, but in the scatter plot's y-axis direction. For low concentration samples, such a view will inflate the proportional distance of an outlying point from the other points. For visualizing anomalous results it may be useful to view both types of difference plots.

9 Quantitative Analysis

Quantitative techniques can be applied to both difference plots and scatter plots to estimate bias.

9.1 Estimating Bias From Difference Plots

When introducing a measurement procedure into a clinical laboratory (where N = 40 is sufficient), only bias estimation from difference plots is required. If further analysis is desired, then the laboratorian should reference Section 9.2, which covers regression techniques of estimating bias. The underlying

assumption of computing bias from difference plots is that some part of the relationship for either constant difference plots or proportional difference plots may have a linear constant bias, either as an absolute difference (constant SD) or as a proportion (constant CV). Therefore, the overall bias estimate can be used for any concentration within that interval. The use of regression analysis, on the other hand, provides a unique estimate of bias at any specified concentration.

9.1.1 Constant Standard Deviation

If, on review of the difference plots, the bias appears to be consistent across the measurement interval on the reporting units difference plot (see Figure 11), then an estimate of the bias between measurement procedures can be made by using the average (or median) of the individual differences between the measurement procedures. This is the bias estimate for any concentration within the measurement interval.

In Figure 11, the vertical distribution is displayed to the right of the difference plot. Note that the histogram to the right shows the typical bell-shaped normal distribution. As an alternative, the mountain plot as described in CLSI document EP21[4] could also be used.

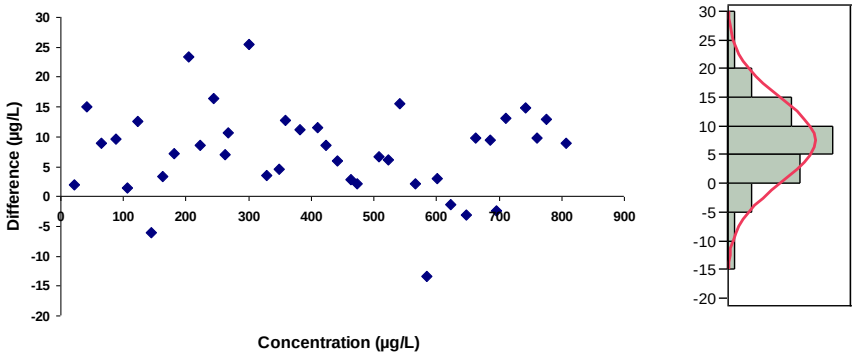

Figure 11. Reporting Units Difference Plot (From Table J1)

For a constant SD difference relationship, the differences to use are provided by the equation $d_i = y_i - x_i$. For a relationship with a nonskewed vertical distribution (as seen in Figure 11), compute bias as the average of all such differences.

$$\bar{d} = \sum_{i=1}^{N} d_i / N \tag{9}$$

For the distribution in Figure 11, this average result is 7.5 µg/L, which is the appropriate estimate for the entire measured interval from 20–800 µg/L.

Figure 12 provides a view of a constant difference relationship with an outlying point that creates a skewed vertical distribution. This outlying point is readily seen in the histogram to the right.

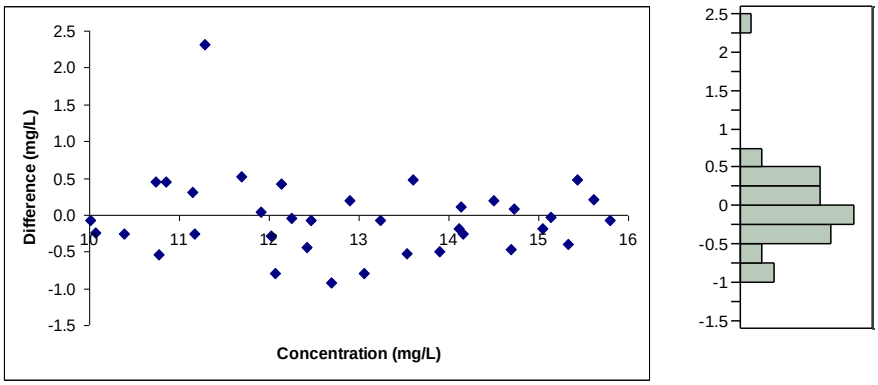

Figure 12. Reporting Units Difference Plot (From Table J5)

For a constant difference relationship with a skewed vertical distribution, compute the bias as the median of the difference values. For the distribution in Figure 12, this median result is −0.07 mg/L, which is the appropriate estimate for the entire measured interval from 10–16 mg/L.

9.1.2 Constant Coefficient of Variation

Figure 13 includes the percent difference plot of Figure 6D, with the horizontal axis being the ranked order number of the samples.

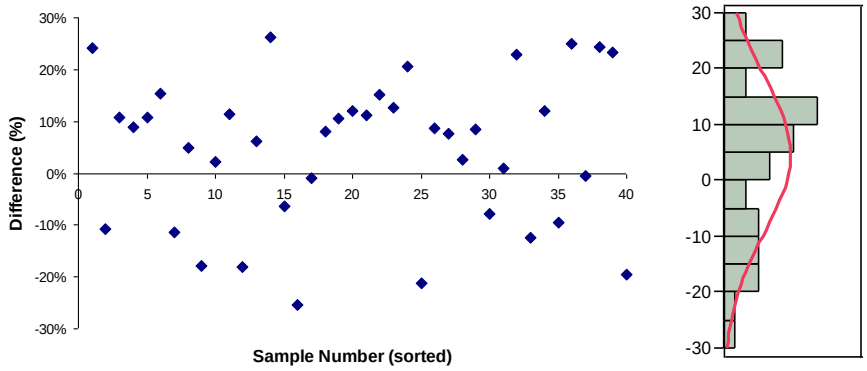

Figure 13. Percent Difference Plot (From Table J3)

Note the histogram on the right of Figure 13. While it does not strictly follow the classical bell-shaped curve it does not display outlying point(s) separated from the primary distribution. Therefore, for this constant CV distribution, no outlying percent differences significantly skew the results, so the average proportional difference can be used. For this distribution the average is 4.6%. In other instances, where significant skewness is detected, the median can be used. Note that once the distribution is determined to be proportionally consistent across a range of results, then the bias is calculated only in the vertical direction, and the horizontal axis is irrelevant. This means that both the average bias across the measuring interval and the bias at any specified concentration is estimated by this same calculated bias.

For the constant SD calculations above, the procedures for calculating both the horizontal and the vertical axis results are irrelevant to the calculation of bias. For constant CV calculations, however, the vertical difference axis is a ratio of difference to concentration. In this case, the user must decide between equations (2) and (4) in Table 3; in other words, dividing by the comparative measurement procedure or by the average of the candidate and comparative measurement procedures. The same difference calculations for *d* (vertical axis) are presented as equations (6) and (8) in Table 4 as seen in Figure 13.

In Figure 14, the constant CV data from Figure 10C are presented where the difference is divided by the comparative measurement procedure.

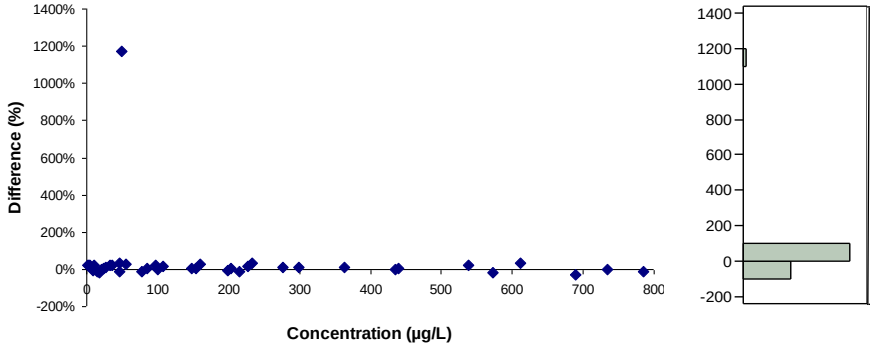

Figure 14. Percent Difference Plot (From Table J4)

In this case, the median result is 7.5% and the average result is 36.5%. Clearly, the median result is the best estimate of central tendency in this case, but the difference plot is too compressed to tell the user whether this estimate is usable over the interval of the measurements. Appendix B presents a method to determine if a result can statistically be declared an outlier. When an outlier is identified (see Figure 14) the data can additionally be presented without this outlier, as seen in Figure 15.

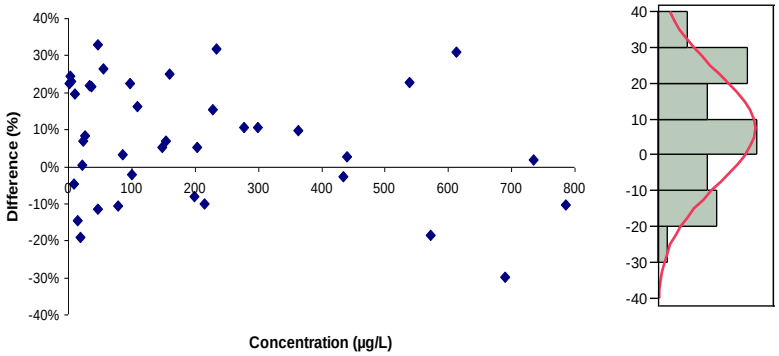

Figure 15. Percent Difference Plot Without Outlier (From Table J4)

Figure 15 shows consistent proportional differences across the measured interval of this study. Therefore, a single bias estimate can reasonably be used to represent the bias within this interval from approximately 0–800 µg/L. This bias estimate is the median of all sample results, including the outlier, which was stated above as 7.5%. (Note that in this case, the average result after excluding the outlier is also 7.5%.)

9.1.3 Mixed Variability (Standard Deviation and Coefficient of Variation)

If a mixed variability model is observed, determine the sample rank, *k*, that best separates the constant SD portion of the dataset from the constant CV portion (see Figure 7). For the low concentration portion, compute the estimate of bias as a constant difference relationship over the samples having rank 1 to *k*. For the high concentration portion, compute the estimate of percent bias as a proportional difference relationship over the samples having rank $k+1$ to N.

Use at least 20 samples in each subgroup for a reasonably accurate estimate of bias for both intervals. After computing the bias (or percent bias) of each subgroup, report this estimate plus the applicable concentration interval of each sample subgroup. The data from Figures 5 and 7 were analyzed in this way. The lowest 40 points (from 0–1.8 µg/L) had an average offset of 0.20 µg/L. The highest 39 points (from

1.8–96 µg/L) had an average proportional offset of 0.43%. See Appendix I to see how these results were computed.

When the precision profiles of measurement procedures have been characterized, the expected behavior of a comparison study, including, for example, a constant SD at low concentrations and a constant CV at high concentrations, can be provided as acceptance criteria. An example of this type of analysis is included in Appendix D.

9.1.4 When Bias Changes With Concentration

If the bias changes over the measurement interval in a linear fashion, then the dataset is inappropriate for the bias estimation techniques described in this section. In these instances the user should perform a regression analysis, as described in Section 9.2.

To determine if such a change over the measuring interval is significant, a regression can be performed on the difference plot. If the contribution of slope is significantly different from zero with an approximate 95% level of confidence, then a nonconstant difference is present, as seen in Figure 16. An ordinary linear regression (OLR) was performed to determine the line fit in Figure 16 because, in this case, the variability around the line is relatively uniform across the concentration interval of the data. If this assumption is not met, other regression techniques[26] (beyond the scope of this guideline) may be more appropriate.

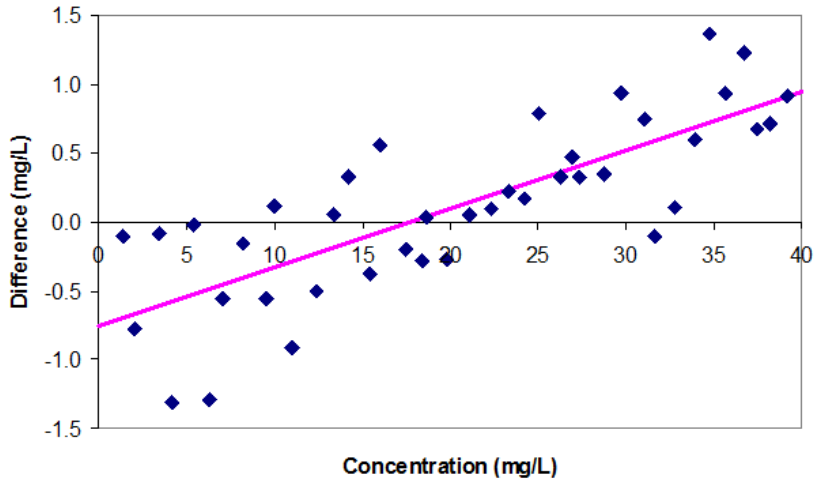

Figure 16. Regression Fit of Bias vs Concentration (Same Data as Figure 8)

9.1.5 Nonlinear Relationship

If the bias displays a nonlinear relationship with concentration, as seen in Figure 9, then neither difference plots nor regressions over the entire measuring interval are appropriate. One option for such data is an analysis that describes fitting a line (perhaps nonlinear) to the points provided in a difference plot, as described by Hawkins.[27] Such analysis is beyond the scope of this guideline.

If the goal of the study is to provide a bias estimate for a specific medical decision concentration X_c, then the points around that concentration can be used to provide such an estimate. A difference plot of a subset of points around the medical decision concentration can be constructed. At least 20 such points should be selected by either selecting the nearest 10 points above and 10 points below the concentration or selecting an interval of concentrations around the medical decision concentration. The above selection of points should be performed based on a list of results ranked on the average of the candidate and comparative measurement procedures (ie, equations 7 or 8 in Table 4). Selecting points for computing such an estimate based on a list ranked on a single measurement procedure would improperly bias the results.[28]

9.1.6 Vertically Skewed Distribution (Aberrant Results)

If one sample or a small number of samples is causing the perceived skewness in the y-axis direction (see Figure 10), use the techniques outlined in Appendix B to determine if they are true outliers. If so, investigate the possible cause of the outlying result. If the point(s) cannot be eliminated for cause, then the vertical distribution will remain nonsymmetrical and the bias should be estimated using the median difference or median percent difference.

9.1.7 Confidence Interval of the Bias Estimate

Once the average (or median) bias and its CI have been determined, they can also be used to evaluate bias at appropriate medical decision concentration based on the acceptance criteria.

9.1.7.1 Bias Estimate

With a symmetrical distribution of differences (either SD or CV), the average is used as the bias estimate. Determining the CI for this estimate requires the computation of its standard error (SE):

$$\mathrm{SE}(\bar{d}) = \sqrt{\frac{\sum_{i=1}^{N}\left(d_i - \bar{d}\right)^2}{N(N-1)}} = SD/\sqrt{N}, \qquad (10)$$

where d_i is the difference between the candidate and comparative measurement procedures for each sample i and $\bar{d}$ is the average of all such differences (see equation 9).

Assuming the differences follow a normal (gaussian) distribution, the CI is computed by multiplying the SE by the factor derived from the confidence desired (typically 95%) and the sample size by using Student t distribution and adding and subtracting the result from the average estimate. For a sample size N of 20 ($N - 1 = 19$ degrees of freedom), this 95% factor is 2.093, for $N = 40$ it is 2.023, and for $N = 100$ it is 1.984. In the first case, the 95% CI would be from (average − 2.093SE) to (average + 2.093SE).

9.1.7.2 Median Bias Estimate

While the CI for an average estimate of bias can be computed by an equation, computing the CI around the estimate of median bias requires a nonparametric method. The interval it produces will be close to but may not be the standard 95%, because it depends upon the number of points used in the estimate. As the number of points decreases, the CI becomes less precise. See Appendix A for an example of computing the CI of a median bias for 100 points.

9.2 Fitting a Line to Scatter Plots (Regression Analysis)

Regression, as an analysis technique, is applicable in a wider range of situations than difference plots. For some, it also provides a more intuitive comparison of each sample point between the two measurement procedures. As with difference plots, regression analysis techniques require that underlying assumptions are met. Difference plots provide many of the answers with respect to such assumptions. Therefore, a good first step in deciding which regression technique to use is reviewing difference plots, as described above, to characterize the distribution of differences.

As mentioned above, studies conducted to introduce a candidate measurement procedure by clinical laboratories may not require analyses beyond difference plot analysis. However, if some assumptions

required by difference plots are not met, then the best choice is to continue on to regression analysis. For manufacturers conducting establishment or validation studies, regression analysis is necessary.

The initial goal of a regression analysis in a measurement procedure comparison study is to fit a straight line though the data presented as an scatter plot with the comparative measurement procedure on the x-axis and the candidate measurement procedure on the y-axis. The default assumption is trying to demonstrate exact concordance (identity) between both measurement procedures. In other words, if the comparative measurement procedure provides a result of 1 for a sample (x_1), then the candidate measurement procedure will also provide 1 (y_1), and if the comparative measurement procedure provides 100 for another sample (x_2), the candidate measurement procedure will also provide 100 (y_2).

It is easy to draw a line, defined as $y = a + bx$, through two such perfect results, where x is the comparative measurement procedure result, y is the candidate measurement procedure result, a is the intercept of the line to the y-axis, and b is the slope of the line. In this case, the slope would be $(y_1 - y_2) / (x_1 - x_2) = (1 - 100) / (100 - 1) = 1.0$. The intercept would then be $a = y_2 - bx_2 = 100 - 1.0 \bullet 100 = 0.0$. Therefore, the equation for a measurement procedure comparison between two perfectly concordant measurement procedures has an intercept of zero and a slope of one.

In some cases, it is known beforehand that there will not be such a perfect relationship. One example is a candidate measurement procedure that is trying to more closely match an international standard. The comparative measurement procedure in this case may have been introduced before the standard was developed. Assume that the known positive proportional difference is 20%. In this instance, the expected result of a measurement procedure comparison study is an intercept of zero and a slope of 1.2.

9.2.1 Constant Standard Deviation

Two measurement procedure comparisons are presented below. The first (see Figure 17A), using the data from Figure 1, confirms that there is a consistent amount of variability (SD) across the measuring interval. The amount of scatter around this relationship is small, especially when the entire range of measurements is considered. The same scatter (SD), and relationship of scatter to concentration, is presented in Figure 17B, but the interval over which measurements are made is far more restricted.

The correlation coefficient (r), is often squared to provide the coefficient of determination (r^2). This is the fraction of variance in y explained by a least squares regression line fitted through the data. The theory and calculations for computing r and for fitting a least squares regression line through such data is presented in Appendix C. The computed r^2 for these two examples is 0.999 for Figure 17A and 0.961 for Figure 17B.

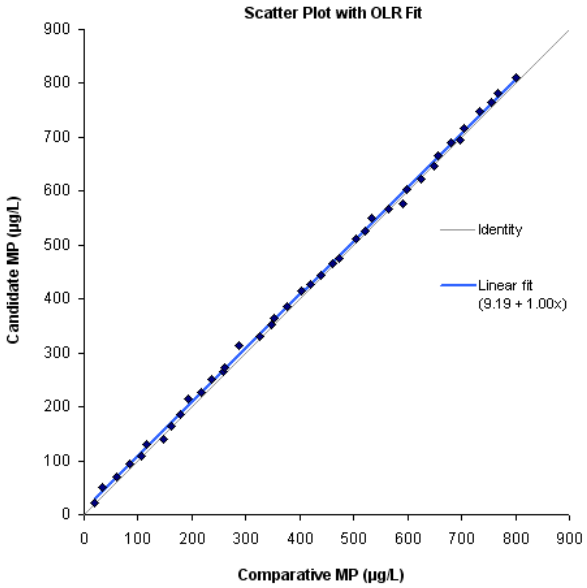

Figure 17A. OLR Fit to Highly Correlated Data

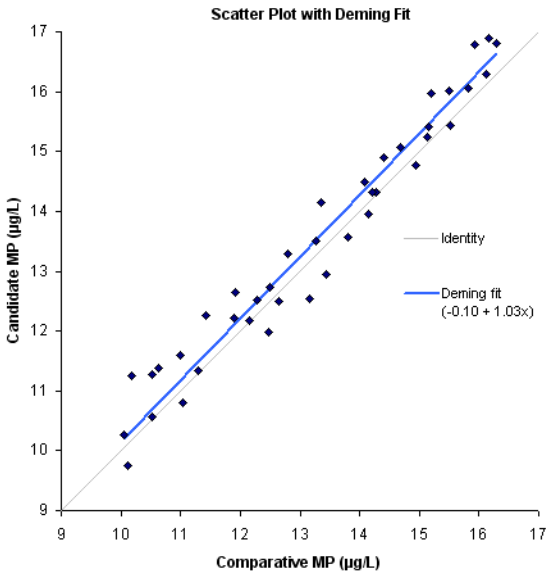

Figure 17B. Deming Fit to Less Correlated Data

Figure 17. Regression Fit to Data With Constant SD (Figure 17A Data From Table J1; Figure 17B Data From Table J6)

Abbreviations: MP, measurement procedure; OLR, ordinary linear regression.

In relationships such as those in Figure 17 in which there is a constant SD, an even distribution of points over the measured interval, and an r^2 greater than 0.95, OLR (see Appendix C) will provide results consistent with other techniques. However, if any of these assumptions are broken, then other techniques, mentioned below, should be used.

In cases of constant difference variability (constant SD), the estimate of r^2 should be used only as an indicator of the strength of a regression fit and as a rule-of-thumb determination of whether an OLR may be acceptable. It should not be used as a criterion to accept or reject the candidate measurement procedure. For multiple replicate situations in which replicate averages are used for each measurement, the correlation coefficient, while an accurate representation of the amount of between-sample variability explained by the regression fit, will overstate the amount of total variability (within and between sample) explained by a regression fit to the individual replicate results. Finally, the correlation coefficient should not be used as an indication of linearity. Follow the procedures in CLSI document EP06[29] to perform such an evaluation.

The above argument for using OLR, for instances similar to those in Figure 17, is basically that it is adequate. A better case can be made for using a constant SD Deming[10] regression for such instances. OLR attempts to minimize the differences between the points and the fitted line as measured in the vertical (y) direction. This technique assumes that only the candidate measurement procedure has inherent imprecision. This is never true; even a comparative measurement procedure comprised of samples made gravimetrically from standard material still has imprecision associated with mass determinations. Deming regression as shown in Figure 17B (see Appendix E) allows the imprecision of both measurement procedures to be taken into account. If the imprecision estimates of both measurement procedures are known, such as from precision studies conducted under CLSI document EP05,[2] then this knowledge can be used to determine the fit.[30]

The consistency of a regression fit using a Deming regression is easily shown by switching the x-axis and the y-axis and redoing the regression. Using Deming regression, the relationship will in most cases be more consistent. Using OLR, the two results will usually be inconsistent.

In conclusion, for cases in which there is constant difference variability (constant SD) across the measurement interval of the two measurement procedures, constant SD Deming is recommended as the default regression technique.

9.2.2 Constant Coefficient of Variation

As seen in Section 8, measurement procedure comparisons with constant difference variability should not use the same difference plot techniques as those with proportional difference variability. This is also true for regression techniques.

Datasets exhibiting proportional difference variability do not meet the underlying assumptions for either OLR or constant SD Deming regression. Instead of a constant SD, such datasets exhibit a constant CV. Much like Figure 2, the scatter plot will display the points opening up like a trumpet with the distribution narrow on the lower left and wider on the upper right (see Figure 18).

Weighted least squares (WLS) regression can take such a distribution into account. The specific weighting known as constant CV least squares regression gives each point a weight inversely proportional to the square of the concentration on the x-axis. Thus, points further to the right have less influence on the regression line fit than do points on the left because they are expected to be more scattered. A discussion of this regression technique is provided in Appendix D.

WLS regression has some of the same faults as OLR. First, the assumption of zero imprecision for the comparative measurement procedure is false. Second, the results are even less likely to be consistent if the axes are switched.

The constant CV Deming regression can solve both of these issues. The specific weighting in constant CV Deming regression, works in a similar fashion as WLS. Points to the upper right have less influence on the regression fit than do points to the lower left. As in the previous discussion of Deming regression, the assumption is that both measurement procedures have some inherent variability, in this case expressed as CV. A regression line fit to the data in Figure 2 using constant CV Deming is shown in Figure 18. A discussion of constant CV Deming regression is provided in Appendix F.

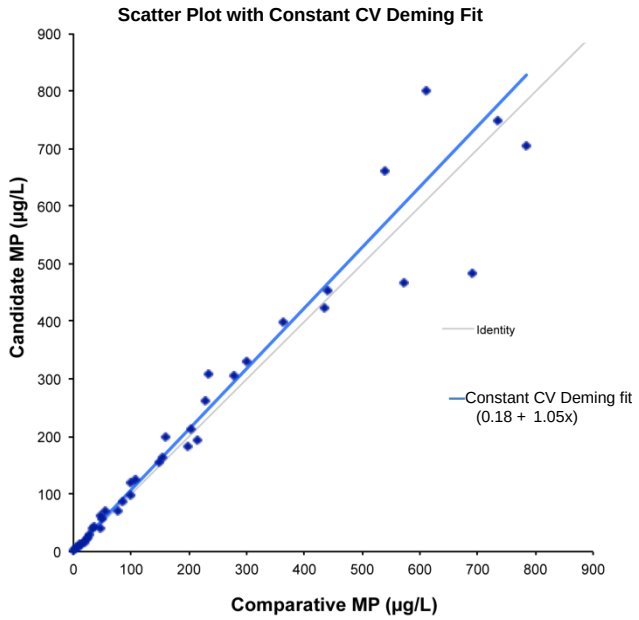

Figure 18. Constant CV Deming Regression Fit to Data With Constant CV (From Table J2)
Abbreviations: CV, coefficient of variation; MP, measurement procedure.

Neither WLS nor constant CV Deming regressions are possible if points with a concentration of zero for either of the two measurement procedures are compared, because weighting of points based on the calculation of CV = SD / concentration is not feasible. However, because the measuring interval is defined as being above the limit of detection or limit of quantitation (see CLSI document EP17)[31] and within the interval of linearity, zero concentration points should not be part of a measurement procedure comparison dataset. In cases of data exhibiting constant CV over the measurement interval, the recommended default regression technique should be constant CV Deming regression.

All examples in this document assume a single estimate of concentration for each measurement procedure for each sample. If this estimate is the average of multiple replicates, then imprecision data are lost that could be used to determine the relative imprecision of the two measurement procedures. The resultant imprecision ratio between the two measurement procedures can be automatically computed by some Deming regression software packages. This ratio, often referred to as lambda, is an input into any Deming regression. If multiple replicates or such software are not available, previously generated imprecision results from studies based on CLSI document EP05[2] can be used to estimate this ratio. If such studies are not available or similar measurement procedures are being compared, the best default estimate for this ratio is 1.0. Without any knowledge of the ratio, for some purposes, it may be desired to vary the ratio to assess the sensitivity of the Deming regression to its value. These considerations hold true for both constant SD Deming regression (see Appendix E) and constant CV Deming regression (see Appendix F).

Some users advocate the transformation of data before plotting constant CV datasets. Both logarithmic and power functions have been used for such transformations. The advantage of such techniques is that the scatter around the line fit can be made to resemble an evenly spaced, constant SD dataset. Such views of the data can help determine whether the spread of points at high concentrations is due solely to a constant CV relationship or to potential outlying points. After transformation, an OLR or constant SD Deming regression fit can usually be performed. The resultant line equation can be used to estimate bias at any concentration after the reverse transformation. However, the slope of the transformed data line fit cannot be used for comparison to bias acceptance criteria, because such criteria are stated in reference to untransformed results.

Finally, for constant CV datasets, many advocate the use of Passing-Bablok regression (see Section 9.2.3). This regression method, as with any nonparametric technique, requires a higher sample size than a parametric technique to reproducibly provide the same results. However, given the suggested sample size for manufacturers (100) and even clinical laboratories (40) it is a viable technique for constant CV datasets.

9.2.3 Mixed Variability (Standard Deviation and Coefficient of Variation)

It is assumed that measurement procedure comparison data have already been explored using difference plots before the initiation of regression analysis. The data in Figure 5B showed, at low concentrations, that the variability of the differences was constant. In Figure 5C the same data showed, at higher concentrations, that there was proportional difference. In such a case, the data display neither a constant SD nor a constant CV over the entire measurement interval. If the data show no significant offset at low concentrations using difference plots (ie, distribution of differences overlap the zero bias line), then the default recommendation of a constant CV Deming regression is a reasonable option. In the case shown in Figure 5, however, the influence of the highest concentration points causes an inflated estimate of proportional difference (see Figure 19A). The influence of each such point would have been less pronounced had more points been collected at higher concentrations, and may have reduced the need to look beyond the constant CV Deming regression technique.

Passing-Bablok regression is a nonparametric technique that, while fitting a line through the data, makes no assumptions about the distribution of the data points. It essentially draws a median line through the data (ie, there will always be close to an equal number of points on either side of a Passing-Bablok fit).

Because it makes no distribution assumptions, it is an appropriate technique to use for datasets, such as those from Figure 5 that break the assumptions made in other techniques. A description of Passing-Bablok regression can be found in Appendix G. The result of a Passing-Bablok regression on the data from Figure 5 is shown in Figure 19B. Again, if this had been a manufacturer's dataset using 100 or more samples, additional samples at higher concentrations should have been collected to more evenly cover the measuring interval.

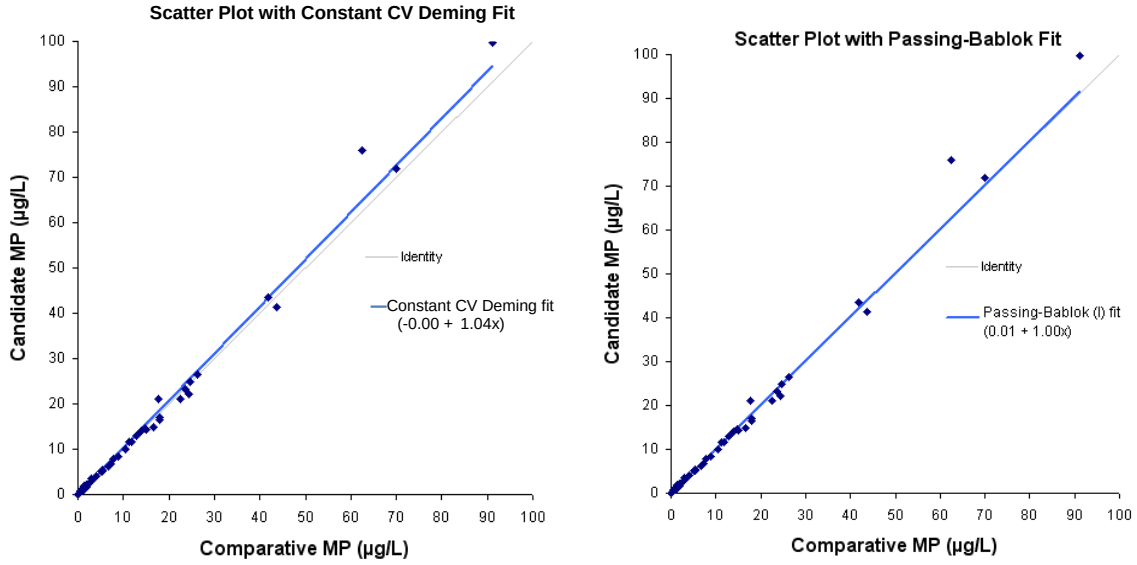

Figure 19A. Constant CV Deming Fit **Figure 19B. Passing-Bablok Fit**

Figure 19. Regression Fits With Mixed Variability Data (From Table I1)
Abbreviations: MP, measurement procedure.

In the earlier difference plot analysis described in Section 9.1.3, the data previously seen in Figure 5 had an average offset at low concentrations of about 0.20 µg/L and an average proportional offset at high concentrations of 0.43%. The Deming regression fit from Figure 19A gives an intercept of 0.00 µg/L and a higher proportional offset of 4.0%. The Passing-Bablok regression fit from Figure 19B more closely matches the difference plot estimates with an intercept of 0.01 µg/L and a proportional offset of 0.3%. See Appendix I for a complete set of results for these data.

In such a case, if the full precision profile of the two measurement procedures is known, this information can be used to weight the Deming regression appropriately across the interval of collected measurements.[32] Some discussion of such an analysis is covered in Appendix F.

In conclusion, in mixed variability cases, constant CV Deming regression is more resistant to the influence of a few high concentration samples than unweighted regression techniques but it will not totally eliminate their influence. In such cases, the Passing-Bablok regression is the better option.

9.2.4 Aberrant Results

Aberrant results create a distribution that is nonsymmetrical, and possibly skewed. Such distributions do not meet the underlying assumptions of either OLR or Deming regression. Therefore, in such situations, a Passing-Bablok regression should be performed.

9.2.5 Nonlinear Distribution

A nonlinear distribution does not meet the underlying assumptions of any of the regression techniques because they all assume linearity and that the two measurement procedures measure the same quantity. In instances where these assumptions are not met, the difference plot techniques described in Section 9.1.5 should be used.

9.3 Bias and Regression Parameters With Confidence Intervals

All regression analyses mentioned above provide an estimate of the relationship between the candidate and comparative measurement procedures by fitting a line through the data with the equation $y = a + bx$ where a is the intercept and b is the slope.

The regression equation provided by any of these methods can be used to estimate the bias between the candidate measurement procedure (Y) and the comparative measurement procedure (X) along the vertical (y-) axis at any value within the interval of measured comparative values.

If an OLR or weighted OLR was performed, the CI (typically 95%) of this estimate can be computed directly. See Appendixes C and D for these computation descriptions.

For all other regression techniques, the CI of any bias estimate cannot be directly computed through an equation. For these techniques, combining the CI of the slope and the CI of the intercept does not directly compute the CI of the bias estimate in the vertical (y-axis) direction at a specified comparative value. Instead, an iterative technique can be used to create a set of data from the N regression points. For each such set created, a regression line can be fit and the bias estimate can be made in the vertical (y-axis) direction. Using these determinations from at least N such datasets, the SE of the bias estimate can be computed. Similarly, the SE of the slope and intercept can also be determined.

One common method for performing this iteration is the jackknife technique, in which each sample is withdrawn in turn from the dataset to create N sample sets of $N - 1$ samples each. This technique is described in detail in Appendix H. A second method is the bootstrap technique in which any number of sample sets are created by sampling randomly with replacement from the original set of samples.[33] In this way, N sample sets of $N - 1$ samples each can also be created with this technique. In doing so, the same parameter estimation equations provided in Appendix H can be used. The use of the jackknife technique is inappropriate for Passing-Bablok regression so the bootstrap technique should be used for this regression method.

10 Comparisons Within a Measurement Procedure

Manufacturers or laboratories may wish to make a comparison of two conditions within an already validated or released measurement procedure. The same analysis methods mentioned earlier in this guideline are applicable. Such studies may be performed to estimate bias across sample tube types, raw materials, reagent lots, calibrator lots, or other factors. The scope of the experimental and data-handling procedures for this purpose will be smaller than that for claims establishment or claims verification studies.

For these types of studies, the assumption is that measurement procedure performance has been established and verified, including bias, imprecision, and linearity across its measuring interval. There is no need to ensure that the full measuring interval is covered by such a study; the only need is to provide a reasonable interval of concentration measurements covering the clinical decision points and both diseased and nondiseased areas of the measuring interval. Because of these considerations, a sample size of 40 is adequate.

In such studies the data are obtained using the same measurement procedure under two conditions; consequently, the comparison and candidate data have very similar performance characteristics except for the condition being examined.

10.1 Sample Type Comparisons

The relationship between measurements from different tube types or sample types, as mentioned in CLSI document I/LA21,[34] is often of interest to manufacturers. Because only a single factor of a measurement procedure is to be characterized, there is no need to sample other factors such as instrument, day, or calibration. Such studies can be conducted on a single instrument, on a single reagent lot, on a single day.

For such studies, aliquots of both sample types collected from a patient can typically be run on the same instrument, within the same timeframe, on the same lot of reagent material, and on the same calibration. If multiple replicates are run for each sample type, the average of the replicates should be used as the individual measurement for each aliquot.

10.2 Other Comparisons

Other study examples include comparisons typically run by a clinical laboratory, such as between lots of reagent material or between instruments of the same or a different manufacturer (ie, production and backup instrument). Unmodified patient samples are easier to collect for such studies than for sample type studies, because only a single sample is required per patient. A reagent lot comparison should be run on the same instrument with both lots run within a short timeframe. An instrument comparison should also be done within a short timeframe, keeping all factors as consistent as possible. Other bias estimation techniques with smaller sample sizes may be used in such instances such as those specified in CLSI document EP31.[35]

11 Interpreting Results and Comparing to Performance Criteria

The difference between a comparative measurement procedure and a candidate replacement measurement procedure is of interest over their entire common measuring interval (usually expressed as slope) or at one or more medically significant concentrations. Compare the CI of the measured bias (average bias over the measuring interval or bias at a specified concentration) (see Section 9.3) with the definition of limits of acceptable bias. Each manufacturer or clinical laboratory should develop its own criteria (in consultation with medical staff and/or the technical literature). These criteria should be predefined before the measurement procedure comparison study, especially if the study will be evaluated by a regulatory authority. These criteria should include the decision rules regarding the possible outcomes of the estimate of bias and its 95% CI as shown in Figure 20. In this figure, each solid dot is an example bias estimate and the vertical line is its 95% CI.

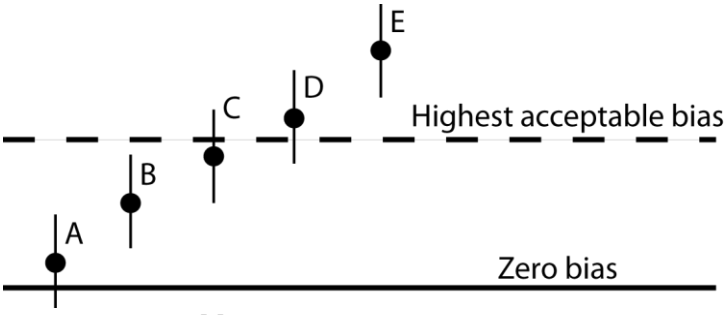

Figure 20. Possible Bias Estimate Outcomes

Outcome A is an instance in which the 95% CI of the measured bias includes zero. The outcome statement can therefore be that no significant bias was seen between the candidate and comparative measurement procedures.

In outcome B the predefined limits of acceptable bias include the 95% CI for estimated bias. Therefore, the bias of the candidate measurement procedure to the comparative measurement procedure meets the bias acceptance criteria with 95% confidence and is acceptable for the defined application. Please note that this statement is also valid for outcome A.

If the estimated bias is within the predefined limits of acceptable bias but the CI is not (outcome C) then it cannot be said that the acceptance criteria was met with 95% confidence. Because the estimated bias is less than the highest acceptable bias limit, it may be concluded that the bias is acceptable. However, an alternate conclusion that an inappropriate percent of sample results will have bias outside the limits is also possible.

If the estimated bias is outside of the predefined limits of acceptable bias but the CI is not (outcome D) then the data do not show a bias outside of the acceptable bias with 95% confidence. Because the CI includes the highest acceptable bias limit, it still may be concluded that the bias is acceptable. However, there is less confidence in this conclusion than for outcome C.

If the estimated bias and its CI are outside of the predefined limits of acceptable bias (outcome E) then the performance of the candidate measurement procedure is not acceptable for the defined application.

Instead of concluding a candidate measurement procedure is not acceptable, the above set of outcomes may instead drive a clinical laboratory to adjust reference intervals using the results of the comparison study per CLSI document EP28.[6] This will not be the case for widely accepted medical decision concentrations. Such cutoff points are established through extensive clinical studies or clinical experience and therefore typically cannot be changed.

Where a manufacturer has provided comparison data for the candidate measurement procedure, the stated bias can replace the acceptance criteria in the analyses above. If the manufacturer's claim for bias is included in the 95% CI (outcomes C or D), then the clinical laboratory can conclude that the candidate measurement procedure meets the bias claims of the manufacturer.

11.1 Manufacturer's Statement of Bias Performance Claims

The following items should be included in a manufacturer's claim for measurement procedure comparison bias. It is expected that manufacturers will provide claims based on regression analysis where *X* is the comparative measurement procedure and *Y* is the candidate measurement procedure.

- The total number of samples (points) used in the measurement procedure comparison. Each sample provides only one point to the comparison.
  - If samples are excluded from the analysis the number of such samples must be stated, along with the reason for their exclusion.

- The interval of collected data (the highest and lowest value of *x* included in the regression).

- The comparative method, and its calibration traceability if known, used in the measurement procedure comparison.

- Whether individual determinations were used in the comparison or averages of replicate determinations and, if so, how many repetitions within each average. This should be noted for both *X* and *Y*.

- The number of days, instruments, reagent lots, calibrator lots, and calibration cycles used to collect the data on *Y*.

- The slope and intercept of the fitted linear regression line (by any method), along with their CIs.

- The bias calculated from the regression line at stated medical decision points (either at generally recognized decision points or at the extremes of the reference interval) along with the CI of each bias estimate.

- A scatter plot of the observed data, using identical scales and intervals for the *x* and *y* axes, with *all* data indicated. The scatter plot should include the fitted regression line with its 95% CI and the line of identity ($y = x$). For display purposes, manufacturers may be required to provide an additional scatter plot of one candidate replicate (x) versus one comparative replicate (Y) if multiple replicates were used.

- The method used to fit the linear regression line (eg, OLR, weighted regression, Deming, Passing-Bablok).

- In cases in which least squares regression is used, the following parameters are to be provided:
  - The SD of residuals from regression $s_{y/x}$ (defined in Appendix C)
  - The correlation coefficient (r) or determination coefficient (r^2)

11.2 Laboratory's Statement of Bias Performance

The laboratory may wish to provide a statement of bias determined through either regression analysis or difference plot analysis. For regression analysis the same, or a subset, of the information listed above for manufacturers can be provided. For difference plot analysis, the details on the measurement procedure comparison study can be the same. Instead of slope and intercept, however, the laboratory should describe 1) the interval over which a constant SD was found and what that bias was in measurement units and 2) the interval over which a constant CV was found and what that bias was in percent difference. The bias at any medical decision point that falls within an interval will be the bias seen within that interval.

References

[1] CLSI. *User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline—Second Edition.* CLSI document EP12-A2. Wayne, PA: Clinical and Laboratory Standards Institute; 2008.

[2] CLSI. *Evaluation of Precision Performance of Quantitative Measurement Methods; Approved Guideline—Second Edition.* CLSI document EP05-A2. Wayne, PA: Clinical and Laboratory Standards Institute; 2004.

[3] CLSI. *User Verification of Performance for Precision and Trueness; Approved Guideline—Second Edition*. CLSI document EP15-A2. Wayne, PA: Clinical and Laboratory Standards Institute; 2006.

[4] CLSI. *Estimation of Total Analytical Error for Clinical Laboratory Methods; Approved Guideline.* CLSI document EP21-A. Wayne, PA: Clinical and Laboratory Standards Institute; 2003.

[5] CLSI. *Interference Testing in Clinical Chemistry; Approved Guideline—Second Edition.* CLSI document EP07-A2. Wayne, PA: Clinical and Laboratory Standards Institute; 2005.

[6] CLSI. *Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline—Third Edition.* CLSI document EP28-A3c. Wayne, PA: Clinical and Laboratory Standards Institute; 2010.

[7] Miller JM, Astles JR, Baszler T, et al.; National Center for Emerging and Zoonotic Infectious Diseases, CDC. Guidelines for safe work practices in human and animal medical diagnostic laboratories. *MMWR Surveill Summ.* 2012;61 Suppl:1-102.

[8] CLSI. *Protection of Laboratory Workers From Occupationally Acquired Infections; Approved Guideline—Third Edition.* CLSI document M29-A3. Wayne, PA: Clinical and Laboratory Standards Institute; 2005.

[9] Bureau International des Poids et Mesures (BIPM). *International Vocabulary of Metrology – Basic and General Concepts and Associated Terms* (VIM, 3rd edition, JCGM 200:2012). http://www.bipm.org/en/publications/guides/vim.html. Accessed July 12, 2013.

[10] Mandel J. *The Statistical Analysis of Experimental Data*. Mineola, New York: Dover Publications; 1984:282-292.

[11] WHO Expert Committee on Biological Standardization. *Glossary of Terms for Biological Substances Used for Texts of the Requirements*. WHO unpublished document BS/95.1793. Geneva, Switzerland: World Health Organization; 1991.

[12] Passing H, Bablok W. A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, Part I. *J Clin Chem Clin Biochem*. 1983;21(11):709-720.

[13] Passing H, Bablok W. Comparison of several regression procedures for method comparison studies and determination of sample sizes. Application of linear regression procedures for method comparison studies in clinical chemistry, Part II. *J Clin Chem Clin Biochem*. 1984;22(6):431-445.

[14] ISO. *Accuracy (trueness and precision) of measurement methods and results – Part 3: Intermediate measures of the precision of a standard measurement method.* ISO 5725-3. Geneva, Switzerland: International Organization for Standardization; 1994.

[15] ISO. *Medical laboratories – Requirements for quality and competence*. ISO 15189. Geneva, Switzerland: International Organization for Standardization; 2012.

[16] ISO. *Accuracy (trueness and precision) of measurement methods and results – Part 1: General principles and definitions.* ISO 5725-1. Geneva, Switzerland: International Organization for Standardization; 1994.

[17] CLSI. *Quality Assurance for Design Control and Implementation of Immunohistochemistry Assays; Approved Guideline—Second Edition.* CLSI document I/LA28-A2. Wayne, PA: Clinical and Laboratory Standards Institute; 2011.

[18] Miller WG, Erek A, Cunningham TD, Oladipo O, Scott MG, Johnson RE. Commutability limitations influence quality control results with different reagent lots. *Clin Chem.* 2011;57(1):76-83.

[19] CLSI. *Evaluation of Matrix Effects; Approved Guideline—Second Edition.* CLSI document EP14-A2. Wayne, PA: Clinical and Laboratory Standards Institute; 2005.

[20] Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. *J Biopharm Stat*. 2007;17(4):571-582.

[21] Carroll RJ, Ruppert D. The use and misuse of orthogonal regression in linear errors-in-variables models. *Am Stat.* 1996; 50(1):1-6.

[22] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307- 310.

[23] Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res.* 1999;8(2):135-160.

[24] Krouwer JS. Why Bland-Altman plots should use X, not (Y+X)/2 when X is a reference method. *Stat Med*. 2008;27(5):778-780.

[25] Bland JM, Altman DG. Comparing methods of measurement: why plotting difference against standard method is misleading. *Lancet*. 1995;346(8982):1085-1087.

[26] Carstensen B. *Comparing Clinical Measurement Methods: A Practical Guide*. Chichester, United Kingdom: John Wily & Sons Ltd; 2010.

[27] Hawkins DM. Diagnostics for conformity of paired quantitative measurements. *Stat Med*. 2002;21(13):1913-1935.

[28] Keller T, Butz H, Lein M, et al. Discordance analysis characteristics as a new method to compare the diagnostic accuracy of tests: example of complexed versus total prostate-specific antigen. *Clin Chem*. 2005;51(3):532-539.

[29] CLSI. *Evaluation of the Linearity of Quantitative Measurement Procedures: A Statistical Approach; Approved Guideline.* CLSI document EP06-A. Wayne, PA: Clinical and Laboratory Standards Institute; 2003.

[30] Linnet K. Performance of Deming regression analysis in case of misspecified analytical error ratio in method comparison studies. *Clin Chem*. 1998;44(5):1024-1031.

[31] CLSI. *Evaluation of Detection Capability for Clinical Laboratory Measurement Procedures; Approved Guideline—Second Edition.* CLSI document EP17-A2. Wayne, PA: Clinical and Laboratory Standards Institute; 2012.

[32] Martin RF. General Deming regression for estimating systematic bias and its confidence interval in method-comparison studies. *Clin Chem.* 2000;46(1):100-104.

[33] Davison AC, Hinkley DV. *Bootstrap Methods and their Application*. Cambridge, UK: Cambridge University Press; 1997.

[34] CLSI. *Clinical Evaluation of Immunoassays; Approved Guideline—Second Edition*. CLSI document I/LA21-A2. Wayne, PA: Clinical and Laboratory Standards Institute; 2008.

[35] CLSI. *Verification of Comparability of Patient Results Within One Health Care System; Approved Guideline (Interim Revision).* CLSI document EP31-A-IR. Wayne, PA: Clinical and Laboratory Standards Institute; 2012.

Appendix A. Confidence Interval of a Median Estimate of Bias Between Measurement Procedures

In instances where the distribution of results does not follow a normal (gaussian) distribution, the median is a more robust estimator of central tendency than the mean. This section provides a procedure to compute the median and its confidence interval (CI). The procedure is based on the Wilcoxon distribution-free signed rank test.

Experiment

Below is the experimental layout of a measurement procedure comparison study, where y_i is the result with the candidate measurement procedure and x_i is the result with the comparative measurement procedure.

Patient	x_i	y_i
1	x_1	y_1
2	x_2	y_2
3	x_3	y_3
…	...	...
N	x_N	x_N

Assumptions

1. Let $d_i = y_i - x_i$, for $i = 1,\ldots N$. The differences $d_1,...d_N$ are mutually independent.
2. Each d_i comes from a continuous population, not necessarily the same, that is symmetrical about a common median θ.

1. Hodges-Lehmann point estimator of θ, $\hat{\theta}$

The Hodges-Lehmann point estimator is given by:

$$\hat{\theta} = median\left(\frac{d_i + d_j}{2}, i \le j = 1,...,N\right) \quad \text{(A1)}$$

Let $W^{(1)} \le \ldots \le W^{(M)}$ denote the ordered values $\frac{d_i + d_j}{2}$, or Walsh average.

From the number of pairs of differences $M = N(N+1)/2$, it follows that:

$$\text{if } M \text{ is odd} \rightarrow k = \frac{M-1}{2}, \text{ then } \hat{\theta} = W^{(k+1)}; \quad \text{(A2)}$$

$$\text{if } M \text{ is even} \rightarrow k = \frac{M}{2}, \text{ then } \hat{\theta} = \frac{W^{(k)} + W^{(k+1)}}{2}. \quad \text{(A3)}$$

Appendix A. (Continued)

2. Tukey two-sided CI for θ, (θ_L, θ_U):

$$\theta_L = W^{(C_\alpha)} \quad \text{(A4)}$$

$$\theta_U = W^{(t_{\alpha/2})} \quad \text{(A5)}$$

$$C_\alpha = \frac{\mathrm{N}(\mathrm{N}+1)}{2} + 1 - t_{\alpha/2} \quad \text{(A6)}$$

The position $t_{\alpha/2}$ for various values of *n* is tabulated in probability tables associated with the Wilcoxon Signed Rank test statistic distribution and is defined as the value under the null distribution, of the Wilcoxon Signed Rank *T* statistic, such that $P(T \geq t) = \alpha/2$. That is, when the probability of a value greater than or equal to $T = \alpha/2$, then $t_{\alpha/2} = T$.

$W^{(1)} \leq \ldots \leq W^{(M)}$ are the ordered valued of the $\frac{d_i + d_j}{2}$ averages, $1 \leq i \leq j \leq \mathrm{N}$, used in computing $\hat{\theta}$. That is, θ_L is the $\frac{d_i + d_j}{2}$ average that occupies the position C_α in the list of *M* ordered $\frac{d_i + d_j}{2}$*s*. And θ_U is the $\frac{d_i + d_j}{2}$ average that occupies the position $t_{\alpha/2}$ in the ordered list.

The calculation of $W^{()}$s is the cumbersome part that would require software for trueness. For example, in an experiment of N = 9, the number of $W^{()}$s, *M*, would be *M* = N(N + 1) / 2 = 45; for an experiment involving N = 50, then *M* = 1275.

For the experiment of N = 50, the upper bound of the 100(1 − α)% CI (α = 0.0495) corresponds to $W^{(t_{0.0495/2})}$, which would be located in position $t_{0.0495/2} = 841$, and the lower bound corresponds to $W^{(C_\alpha)}$, which would be located in $C_{0.0495} = 435$.

The following example, using the data in Table A1 and the plot in Figure A1, demonstrates this computational technique.

Appendix A. (Continued)

Table A1. Median Bias and Confidence Interval Example Data

Patient	*x*	*y*	(*y*−*x*)/*x*		Patient	*x*	*y*	(*y*−*x*)/*x*
1	0.52	0.49	−5.77%		51	27.13	24.42	−9.99%
2	0.99	0.97	−2.02%		52	25.72	27.72	7.78%
3	1.49	0.66	−55.70%		53	24.45	24.17	−1.15%
4	1.99	1.87	−6.03%		54	26.39	27.65	4.77%
5	2.52	2.56	1.59%		55	28.76	28.12	−2.23%
6	3.20	2.91	−9.06%		56	26.37	29.74	12.78%
7	3.37	3.53	4.75%		57	27.63	28.28	2.35%
8	3.95	3.96	0.25%		58	28.78	27.89	−3.09%
9	4.85	4.51	−7.01%		59	27.74	29.19	5.23%
10	4.96	5.46	10.08%		60	32.59	29.34	−9.97%
11	5.33	5.85	9.76%		61	31.48	32.00	1.65%
12	6.00	5.98	−0.33%		62	30.31	31.99	5.54%
13	7.02	6.62	−5.70%		63	32.30	30.50	−5.57%
14	7.15	6.92	−3.22%		64	33.11	32.73	−1.15%
15	7.76	7.47	−3.74%		65	31.33	32.76	4.56%
16	7.71	7.56	−1.95%		66	31.45	31.62	0.54%
17	8.59	8.08	−5.94%		67	31.37	32.35	3.12%
18	9.87	8.94	−9.42%		68	33.54	33.46	−0.24%
19	8.75	9.34	6.74%		69	32.74	32.71	−0.09%
20	10.48	9.81	−6.39%		70	33.21	36.07	8.61%
21	10.16	9.78	−3.74%		71	32.57	32.26	−0.95%
22	11.17	10.91	−2.33%		72	35.85	37.90	5.72%
23	11.83	11.13	−5.92%		73	37.04	34.18	−7.72%
24	11.79	12.17	3.22%		74	35.23	36.43	3.41%
25	12.29	12.63	2.77%		75	34.54	42.28	22.41%
26	11.39	12.96	13.78%		76	35.45	35.52	0.20%
27	13.67	14.93	9.22%		77	39.35	37.29	−5.24%
28	12.93	12.91	−0.15%		78	40.13	39.27	−2.14%
29	12.83	15.35	19.64%		79	37.98	41.93	10.40%
30	16.78	14.71	−12.34%		80	41.87	42.29	1.00%
31	14.72	15.92	8.15%		81	41.14	42.46	3.21%
32	16.53	15.32	−7.32%		82	43.39	37.68	−13.16%
33	17.17	17.04	−0.76%		83	38.93	39.71	2.00%
34	16.82	15.68	−6.78%		84	43.28	42.52	−1.76%
35	18.39	18.17	−1.20%		85	42.48	46.82	10.22%
36	17.68	17.38	−1.70%		86	42.55	44.16	3.78%
37	19.30	18.82	−2.49%		87	45.17	39.29	−13.02%
38	19.53	20.98	7.42%		88	44.18	45.78	3.62%
39	19.77	18.42	−6.83%		89	45.12	41.71	−7.56%
40	20.48	18.77	−8.35%		90	40.93	46.03	12.46%
41	21.08	20.34	−3.51%		91	48.80	46.89	−3.91%
42	21.31	21.37	0.28%		92	49.47	43.86	−11.34%
43	21.64	21.21	−1.99%		93	45.21	47.88	5.91%
44	20.52	24.33	18.57%		94	48.44	46.26	−4.50%
45	24.30	23.68	−2.55%		95	45.07	43.64	−3.17%
46	21.30	21.72	1.97%		96	43.72	47.45	8.53%
47	24.13	23.59	−2.24%		97	49.74	52.83	6.21%
48	23.99	24.19	0.83%		98	47.59	54.06	13.60%
49	22.19	24.19	9.01%		99	48.61	54.09	11.27%
50	23.83	23.75	−0.34%		100	53.08	49.53	−6.69%

Appendix A. (Continued)

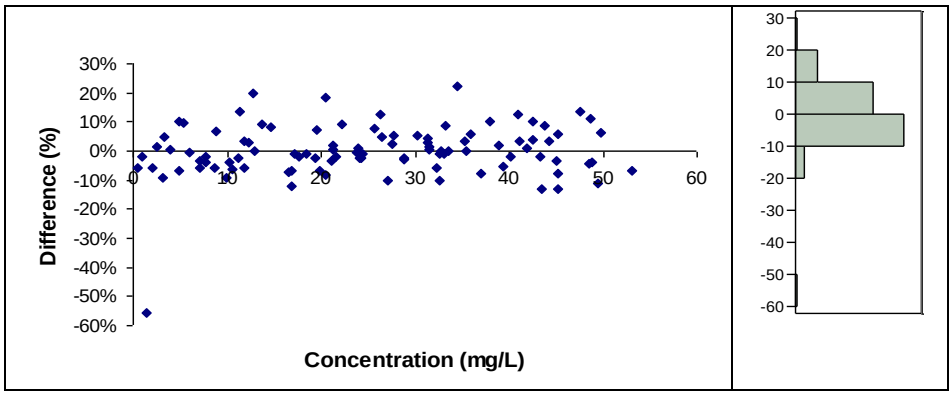

Figure A1. Proportional Difference Plot

N = 100
Median = −0.335%
96.5% CI = −2.020% to 1.590%

References for Appendix A

[1] Hollander M, Wolfe DA. *Nonparametric Statistical Methods*. 2nd ed. New York, NY: John Wiley & Sons, Inc.; 1999.

[2] Harter HL, Owen DB, eds. *Selected Tables in Mathematical Statistics, Volume 1*. Providence, RI: American Mathematical Society; 1973.

Appendix B. Detecting Aberrant Results (Outliers)

The primary goal of detecting aberrant results (outliers) is to enable troubleshooting. They must be found before their underlying cause can be investigated. In an ideal situation, outliers are detected during data collection as suggested in Section 6.7 when their cause can more likely be identified.

This document provides techniques to obtain robust estimates of bias for both difference plots and regression analyses. For difference plots, using the median rather than the average to estimate bias over an interval of measurements reduces or eliminates the undue influence of an outlier on the result. In a similar manner, use of a Passing-Bablok regression provides similar robust regression estimates in the presence of outliers.

The outlier detection process is simplified by using difference plots. Follow the techniques in Section 8.3 of this document to characterize whether the data are from a constant SD or constant CV relationship. In those cases in which there are mixed variability relationships, the data can be split into two sets: concentrations in which there is a constant SD, and concentrations in which there is a constant CV. Within an identified dataset, the differences will be expressed as either differences or percent differences. There should be at least 20 samples in any such set. To align with Section 8.3 of this document, d_j represents a result from the distribution seen in a difference plot.

The detection of aberrant results reduces to a detection of an outlier from within a single distribution. The generalized extreme studentized deviate (ESD) technique, which assumes that the distribution of the vast majority of data points is normal (gaussian), can be used when the number of outliers is unknown, and becomes more robust as the number of samples increases. To perform this technique:

1. Set the significance level (α), which will be used to detect outliers. Typical values are 0.05 or 0.01.

2. Determine if there are potential outliers from graphical or other review of the dataset. No more than 5% of sample results can be flagged as outliers. Set the upper bound on number of potential outliers (h) at this 5% level, rounding down to a whole number. (For 44 samples h will equal 2. For 112 samples, h will equal 5.)

3. For each dataset, determine if one or more suspect results can be statistically deemed outliers based upon the generalized ESD test.[1-3] If the results are not determined to be outliers, then they should be retained in the dataset.

    a. Compute the average ($\bar{d}$) and *SD*, including the suspected outliers.
    b. Find the maximum observed deviate from the average deviate scaled in terms of *SD* (for $j = 1, 2, \ldots N$):

$$ESD_1 = \max\left(\left|d_j - \bar{d}\right|\right) / SD. \quad \text{(B1)}$$

    Repeat this calculation to obtain the (ESD_i) for all potential outliers for $i = 1, 2, \ldots h$. Each subsequent calculation of ESD_i is performed after removing the previously identified potential outlier from the dataset. Thus at each iteration the number of results is reduced by one, then $\bar{d}$, *SD*, and the (ESD_i) are computed again (ie, to look for outlier 2, the number of samples remaining is $j = 1, 2, \ldots N-1$).

Appendix B. (Continued)

c. Corresponding to the number of test statistics (h), compute the following h critical values:

$$\lambda_i = \frac{t_{\nu,p}(\mathrm{N}-1)}{\sqrt{(\mathrm{N}-i+1)(\nu+t_{\nu,p}^2)}}, \tag{B2}$$

where N is the initial number of samples in the dataset, and $i = 1, 2, \ldots h$,

$$\nu = \mathrm{N}-i-1, \tag{B3}$$

$$p = \frac{\alpha}{2(\mathrm{N}-i+1)} \tag{B4}$$

and $t_{\nu,p}$ is the 100p percentage point from Student t distribution with ν degrees of freedom and probability = p.

d. The number of outliers is determined by finding the largest i such that $ESD_i > \lambda_i$.

When an outlier is detected, part of the investigation should be separating out the individual replicates for that sample to determine if the aberrant result is due to a single replicate or not. This may be a key finding that can help determine the cause of the aberrant result.

The dataset from Appendix A (Table A1) is an example of an instance in which an obvious outlier would warrant the use of the generalized ESD test. This dataset provides the following results.

Setting $\alpha = 0.01$, with N = 100 and then $h = 5$. Table B1 lists each subsequent iteration:

Table B1. Example Results

Parameter	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5 = h$
Average $(\bar{x})$	0.01%	0.57%	0.35%	0.15%	−0.03%
SD	9.15%	7.25%	6.94%	6.69%	6.45%
ESD_i	6.09	3.01	2.78	2.75	2.14
λ_i	3.90	3.90	3.89	3.89	3.89
Bias	−55.70% $j = 3$	22.41% $j = 75$	19.64% $j = 29$	18.57% $j = 44$	13.78% $j = 28$

Definitions: *ESD*, extreme studentized deviate; λ_i, critical value; j, the row in Table A1 in Appendix A where each bias was obtained; *SD*, standard deviation.

In this series of calculations the only case in which $ESD_i > \lambda_i$ is the first iteration. This is therefore the only identified outlier.

References for Appendix B

1 Rosner B. Percentage points for a generalized ESD many-outlier procedure. *Technometrics.* 1983;25(2):165-172.

2 NIST. Section 1.3.5.17: Detection of outliers. In: *NIST/SEMATECH e-Handbook of Statistical Methods.* http://www.itl.nist.gov/div898/handbook/. Accessed July 12, 2013.

Appendix B. (Continued)

[3] Hypothesis testing: two-sample inference: In: Rosner B. *Fundamentals of Biostatistics*. 5th ed. Pacific Grove, CA: Duxbury; 1999:300-306.

Appendix C. Ordinary Linear Regression

C1 Correlation

The results of an ordinary linear regression (OLR) analysis are valid only if certain assumptions about the data are true. One of these assumptions is that the *X* variable is known without error. In the clinical laboratory, this is not true because every measurement has intrinsic error. However, if the range of measured concentrations is sufficiently wide, the effect of this error on the regression estimates can be considered negligibly small. The correlation coefficient, *r*, can be used as a rough guide to assess the adequacy of the *X* range in overcoming this problem. The formula for r is as follows:

$$r = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}} \quad \text{(C1)}$$

Where x_i is the best estimate of the concentration of measurand in sample "*i*" from the comparative measurement procedure (average over all replicates from that sample), y_i is a similar estimate using the candidate measurement procedure (average over all replicates for that sample), and

$$\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N} \quad \text{(C2)}$$

$$\bar{y} = \frac{\sum_{i=1}^{N} y_i}{N} \quad \text{(C3)}$$

Using the underlying assumptions of OLR, a practical rule has been that the range of *X* can be considered adequate if $r \geq 0.975$ (or, equivalently, if $r^2 \geq 0.95$). Under these assumptions, an *r* that satisfies this requirement indicates that the error in *X* is adequately compensated by the range of data, and OLR can be used to estimate the slope and intercept. If the data do not fit the assumptions of OLR, then this practice is not valid.

NOTE: This procedure assesses the *range* of the data; it does not measure the *distribution* of the data within the measurement interval. One must still obtain an even distribution of data throughout the measurement interval.

C2 Regression Fit

For the set of paired observations (x_i, y_i), the slope (*b*) and the y-intercept (*a*) are calculated according to the following formulas:

Appendix C. (Continued)

The regression parameters are:

$$b = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{N}(x_i - \bar{x})^2} \quad \text{(C4)}$$

$$a = \bar{y} - b\bar{x} \quad \text{(C5)}$$

Thus, the computed line is described by the following equation:

$$\hat{y}_i = a + bx_i \quad \text{(C6)}$$

C3 Residuals

The difference, measured in the *Y* direction, between a given data point and the regression line is called the *residual* for that point. The SD from regression s_{yx} is the SD of these residuals and is thus a measure of the "scatter" of the points around the regression line. The residual for a point (x_i, y_i) can be calculated using the following formula:

$$\text{Residual}_i = y_i - \hat{Y}_i = y_i - (a + bx_i) \quad \text{(C7)}$$

The SD from regression is given by performing the following calculations:

$$s_{yx} = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N-2}} \quad \text{(C8)}$$

The estimate of the predicted bias (B_c) at a given medical decision level, X_c, is given by:

$$\hat{B}_c = a + (b - 1)X_c \quad \text{(C9)}$$

The 95% confidence interval for B_c (the true bias at X_c) is given by:

$$\left[\hat{B}_{c,low}, \hat{B}_{c,high}\right] = \hat{B}_c \pm t(N-2, 0.975)s_{yx}\sqrt{\frac{1}{N} + \frac{(X_c - \bar{x})^2}{\sum_{i=1}^{N}(x_i - \bar{x})^2}} \quad \text{(C10)}$$

where *t* is the Student *t* distribution.

Appendix D. Weighted Least Squares Regression (Weighted Ordinary Linear Regression)

Constant SD assumption is generally unrealistic for most of the clinical laboratory measurement procedure comparisons. The implementation of ordinary linear regression (OLR) in these cases may not be appropriate. The presence of unequal SD is evidenced by the inspection of the difference plots, as explained in Section 8.2 of this document. The inspection of residual plots after fitting an OLR can also indicate the presence of unequal SD.

This approach is called "weighted" because it introduces weights that are inversely related to the square of the SD at a particular concentration as:

$$w_i = \frac{1}{\sigma^2{}_i} \tag{D1}$$

where σ_i is the SD at that particular concentration.

The SD is often assumed to be proportional to the concentration of the candidate measurement procedure. If such an assumption is made, then the weights can be directly determined from the square of this concentration. Alternatively, if a precision profile of the candidate measurement procedure is available, then the SD of each sample and, thus, the weight for each sample can be computed from this profile.

However, SD may or may not be proportional to the concentration, and precision profiles may or may not be available. Repeated measurements on each sample can provide estimates of repeatability at each concentration, but the number of replicates should be large enough to obtain realistic estimates of σ_i.

When information about σ_i is not known, w_i needs to be estimated from the data. The approach described by Neter et al.[1] is presented below.

First, calculate slope (b) and intercept (a) using OLR. Then, calculate residuals as:

$$e_i = Y_i - \hat{Y}_i = Y_i - (a + bX_i) \tag{D2}$$

The absolute value of the residuals is an estimate of SD, as $\sigma_i = |e_i|$. SD function is calculated by regressing σ_i to X_i using OLR. Assuming proportional relationship between SD and concentration, the linear equation is expressed as:

$$\hat{\sigma}_i = a_\sigma + b_\sigma X_i \tag{D3}$$

where a_σ and b_σ are intercept and slope. Use the fitted values from the SD function to estimate the weight as:

$$\hat{w}_i = \frac{1}{\hat{\sigma}^2{}_i} \tag{D4}$$

Appendix D. (Continued)

Calculate weighted average of reference and candidate measurement procedures as:

$$\bar{X}_w = \frac{\sum_{i=1}^{N} x_i\, w_i}{\sum_{i=1}^{N} w_i}, \qquad \bar{Y}_w = \frac{\sum_{i=1}^{N} y_i\, w_i}{\sum_{i=1}^{N} w_i} \tag{D5, D6}$$

Slope and intercept are calculated as:

$$b_w = \frac{\sum_{i=1}^{N} w_i x_i y_i - \frac{\sum_{i=1}^{N} w_i x_i \sum_{i=1}^{N} w_i y_i}{\sum_{i=1}^{N} w_i}}{\sum_{i=1}^{N} w_i x_i^2 - \frac{(\sum_{i=1}^{N} w_i x_i)^2}{\sum_{i=1}^{N} w_i}}, \quad a_w = \bar{Y}_w - b_w \bar{X}_w \tag{D7, D8}$$

Repeat this process several times by using the residuals from the weighted least squares (WLS) fit to re-estimate the SD function, and then obtain revised weights. Usually, after three to four iterations the differences between consecutive estimates of slope and intercept are negligible.

The denominator in the formula of the weighted slope is the weighted sum of squares of X:

$$SSXw = \sum_{i=1}^{N} w_i x_i^2 - \frac{(\sum_{i=1}^{N} w_i\, x_i)^2}{\sum_{i=1}^{N} w_i} \tag{D9}$$

Residuals are calculated using the weighted slope and intercept as:

$$e_i = Y_i - (a_w + b_w x_i) \tag{D10}$$

SD of regression is:

$$s_{yx} = \sqrt{\frac{\sum_{i=1}^{N} e_i^2\, w_i}{N-2}} \tag{D11}$$

Appendix D. (Continued)

Standard errors (SEs) of slope and intercept are:

$$\hat{\sigma}_b = \frac{s_{yx}}{SSXw}, \qquad \hat{\sigma}_a = s_{yx}\sqrt{\frac{1}{\sum_{i=1}^{N} w_i} + \frac{\bar{X}_w^2}{SSXw}} \qquad \text{(D12, D13)}$$

The $100(1-\gamma)\%$ confidence intervals (CIs) for the slope and intercept are:

$$a_w \pm t(N-2,1-\gamma/2)\hat{\sigma}_a \qquad \text{(D14)}$$

$$b_w \pm t(N-2,1-\gamma/2)\hat{\sigma}_b \qquad \text{(D15)}$$

where $t(N-2,1-\gamma/2)$ is the $100(1-\gamma)$ percentile of the *t* distribution with $N-2$ degrees of freedom. Bias at medical decision level X_c is calculated as:

$$\hat{B}_c = a_w + (b_w - 1)X_c \qquad \text{(D16)}$$

SE of bias is:

$$\hat{\sigma}_{Bias} = s_{yx} = \sqrt{\frac{1}{\sum_{i=1}^{N} w_i} + \frac{(X_c - \bar{X}_w)^2}{SSXw}} \qquad \text{(D17)}$$

Assuming the Y_is follow a normal (gaussian) distribution, the $100(1-\gamma)\%$ CI for bias is:

$$\hat{B}_c \pm t(N-2,1-\gamma/2)\hat{\sigma}_{Bias} \qquad \text{(D18)}$$

Data analysis for comparing platelets on two analyzers is shown for illustration purposes. The shape of the scatter of the difference plot shown in Figure D1 indicates that SD is not constant throughout the measurement interval. The data for this example can be found in Table D2.

Appendix D. (Continued)

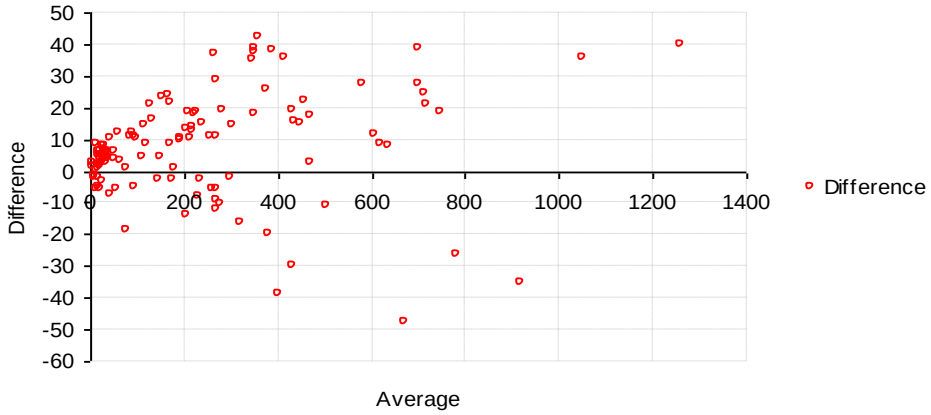

Figure D1. Difference Plots for Comparing Platelets on Two Analyzers

The estimates of the regression parameter using the WLS approach are shown in Table D1.

Table D1. The Estimates of Regression

	Estimate	SE	Lower	Upper
Intercept	3.013	1.073	0.889	5.138
Slope	1.021	0.007	1.007	1.035
Regression SE	1.222			

Abbreviation: SE, standard error.

It is expected that bias throughout the measurement interval will be within ± 10 cells/µL or 5% of the values of the reference measurement procedure. The estimated bias, 95% confidence limits, and specifications are graphically shown in Figure D2. Bias and concentration are shown on original scale as cells per microliter.

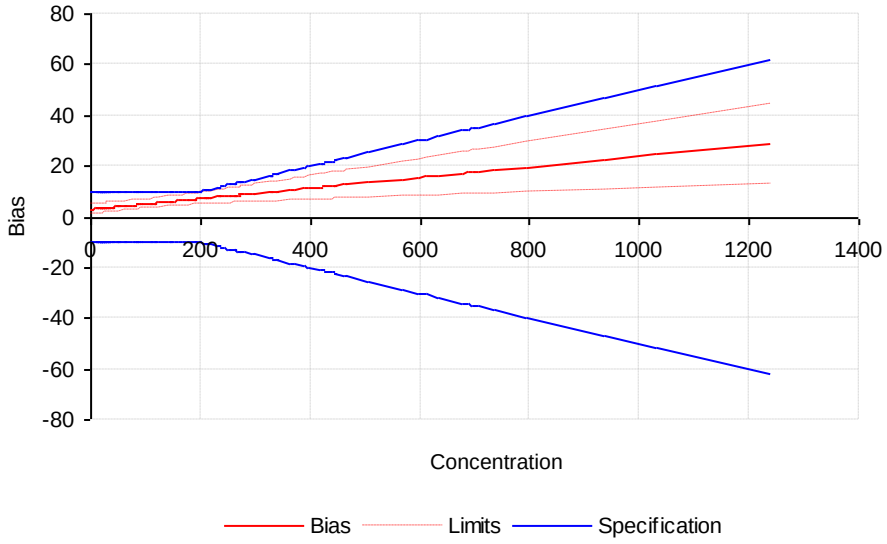

Figure D2. Estimated Bias, Confidence Limits, and Specifications

Appendix D. (Continued)

Table D2. Data for Comparing Platelets on Two Analyzers

Sample	Reference	Test		Sample	Reference	Test		Sample	Reference	Test
1	1.5	3.0		41	60.2	54.8		81	270.1	289.2
2	4.0	6.9		42	61.5	64.6		82	271.3	265.7
3	9.2	8.0		43	78.0	78.6		83	273.5	264.5
4	10.2	18.5		44	80.6	91.4		84	274.2	262.2
5	11.2	9.0		45	84.4	65.7		85	281.1	271.1
6	12.4	13.0		46	85.3	97.2		86	297.0	311.7
7	14.8	19.7		47	89.0	100.0		87	298.7	296.5
8	14.8	16.0		48	92.6	103.2		88	326.7	310.2
9	15.9	21.9		49	94.9	89.6		89	327.1	362.1
10	16.4	10.8		50	108.6	123.4		90	329.6	368.5
11	17.6	22.6		51	110.4	115.0		91	332.8	370.6
12	18.1	15.9		52	115.6	124.4		92	337.4	379.5
13	18.1	20.0		53	116.9	138.1		93	340.1	358.3
14	19.2	14.0		54	122.7	139.2		94	364.8	390.6
15	19.6	25.9		55	143.6	166.8		95	370.1	408.4
16	19.9	21.8		56	146.1	143.7		96	390.6	371.0
17	20.4	24.5		57	146.2	150.8		97	395.7	431.7
18	21.2	29.2		58	154.5	178.5		98	419.3	438.7
19	22.0	27.0		59	161.7	183.4		99	421.3	382.3
20	22.2	24.0		60	167.7	176.1		100	426.3	441.8
21	23.4	25.8		61	176.6	173.7		101	440.4	455.6
22	25.2	22.0		62	179.7	180.4		102	443.4	465.8
23	25.5	19.7		63	188.9	198.9		103	446.2	416.4
24	25.6	33.4		64	189.0	199.4		104	462.7	480.3
25	26.3	30.0		65	197.9	211.1		105	467.7	470.7
26	26.4	28.9		66	201.7	220.1		106	507.4	496.7
27	27.5	34.3		67	207.7	218.3		107	568.3	595.9
28	28.2	34.3		68	209.2	223.4		108	599.6	611.0
29	30.3	35.8		69	210.5	196.8		109	613.8	622.3
30	31.4	37.8		70	210.9	223.8		110	633.5	641.3
31	32.9	37.1		71	214.1	232.2		111	678.6	717.5
32	33.9	40.3		72	218.6	237.1		112	687.6	714.9
33	34.3	37.1		73	232.9	247.9		113	695.1	647.3
34	35.3	40.0		74	235.0	227.0		114	701.0	725.6
35	38.4	42.2		75	237.8	235.3		115	708.3	729.5
36	39.2	49.3		76	246.1	283.0		116	735.6	754.5
37	48.2	41.0		77	252.6	263.5		117	794.8	768.5
38	49.0	55.0		78	254.9	283.5		118	937.0	901.6
39	51.3	55.0		79	261.4	272.3		119	1031.9	1068.0
40	52.2	64.6		80	262.4	256.6		120	1239.3	1279.0

Appendix D. (Continued)

Reference for Appendix D

[1] Neter J, Kutner MH, Wasserman W, Nachtsheim CJ. *Applied Linear Statistical Models*. Chicago, IL: McGraw-Hill/Irwin; 1996.

Appendix E. Deming Regression

With measurement errors in both measurement procedures being compared, the linear model describing the relationship between the measurement procedure results *X* and *Y* can be expressed as:

$$Y = a + b(X + \varepsilon_X) + \varepsilon_Y \quad \text{(E1)}$$

where

a, b = intercept and slope of the linear model, and

$\varepsilon_X, \varepsilon_Y$ = random errors in the *X* and *Y* measurement procedures.

Equation (E1) parameters (a, b) can be estimated with data using regular Deming regression under the following assumptions: the random errors $\varepsilon_X, \varepsilon_Y$ are independent (across the measurement procedures, specimens, and replicates) and normally distributed with zero averages and constant, measurand-level–independent standard deviations (SDs), $\sigma(\varepsilon_X), \sigma(\varepsilon_Y)$.

The SDs, $\sigma(\varepsilon_X), \sigma(\varepsilon_Y)$, of the random errors are practically constant for measurement procedures with small analytical measurement intervals of the measurand, such as electrolytes. In other cases, the SDs of random measurement errors are often approximately proportional to the measurand level over a large proportion of the measuring interval. In such cases, constant CV Deming analysis is more appropriate, as described in Appendix F.

The information on the SDs of the random errors of measurements' approximate constancy or proportionality to the measurand level is often available from the manufacturers' specifications. If such information is not available, it can be obtained by calculating the SDs of the replicated results of measurements for the tested samples, plotting those vs respective replicate averages, and visually examining the graph. The SDs, $\hat{\sigma}(\varepsilon_{X_i}), \hat{\sigma}(\varepsilon_{Y_i})$, of the replicate measurements are calculated for each sample using the following equations (assuming the same number of replicates, *R*, for each measurement procedure, *X* and *Y*, and each of the N samples tested):

$$\hat{\sigma}(\varepsilon_{X_i}) = \sqrt{\frac{1}{R-1}\sum_{j=1}^{R}(X_{ij} - \bar{X}_i)^2} \quad \text{(E2)}$$

$$\hat{\sigma}(\varepsilon_{Y_i}) = \sqrt{\frac{1}{R-1}\sum_{k=1}^{R}(Y_{ik} - \bar{Y}_i)^2} \quad \text{(E3)}$$

i = sample number; $i = 1, 2, \ldots, N$, and

j, k = replicate number; $j, k = 1, 2, \ldots, R$

Appendix E. (Continued)

In the calculations below, $\overline{X}_i, \overline{Y}_i$ are the replicate averages when the SDs of the random measurement errors are approximately constant, and they are the averages of the logarithms of replicate measurement results when the SDs are approximately proportional to the measurand level. The replicate averages for these measurement results obtained with the *i*-th sample ($i = 1, 2, \ldots, N$) are calculated as:

$$\overline{X}_i = \frac{1}{R}\sum_{j=1}^{R} X_{ij} \tag{E4}$$

$$\overline{Y}_i = \frac{1}{R}\sum_{k=1}^{R} Y_{ik} \tag{E5}$$

Equation (E1) can be rewritten for the replicate averages as:

$$\overline{Y}_i = a + b\left(\overline{X}_i + \varepsilon_X\right) + \varepsilon_Y \tag{E6}$$

where $\varepsilon_{\overline{X}}, \varepsilon_{\overline{Y}}$ are random errors of the replicate averages.

Regular Deming regression provides unbiased minimum variance estimates of the equation E6 parameters (a, b) [1] with modified notation; equation for b assumes positive $\hat{\sigma}_{\overline{XY}}$, which is the case with clinical laboratory measurement procedures:

$$b = \frac{\hat{\sigma}^2_{\overline{Y}} - \hat{\lambda}\hat{\sigma}^2_{\overline{X}} + \sqrt{\left(\hat{\sigma}^2_{\overline{Y}} - \hat{\lambda}\hat{\sigma}^2_{\overline{X}}\right)^2 + 4\hat{\lambda}\hat{\sigma}^2_{\overline{XY}}}}{2\hat{\sigma}_{\overline{XY}}} \tag{E7}$$

$$a = \overline{\overline{Y}} - b\overline{\overline{X}} \tag{E8}$$

The parameters used in equations (E7) and (E8) are calculated using formulas (E9) to (E14).

$$\hat{\sigma}^2_{\overline{X}} = \frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}}\left(\overline{X}_i - \overline{\overline{X}}\right)^2 \tag{E9}$$

$$\hat{\sigma}^2_{\overline{Y}} = \frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}}\left(\overline{Y}_i - \overline{\overline{Y}}\right)^2 \tag{E10}$$

$$\hat{\sigma}_{\overline{XY}} = \frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}}\left(\overline{X}_i - \overline{\overline{X}}\right)\left(\overline{Y}_i - \overline{\overline{Y}}\right) \tag{E11}$$

$$\overline{\overline{X}} = \frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}}\overline{X}_i \tag{E12}$$

$$\overline{\overline{Y}} = \frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}}\overline{Y}_i \tag{E13}$$

$$\hat{\lambda} = \hat{\sigma}^2\left(\varepsilon_{\overline{Y}}\right)/\hat{\sigma}^2\left(\varepsilon_{\overline{X}}\right) \tag{E14}$$

where

Appendix E. (Continued)

N	=	number of samples used for fitting model (E3).
$\bar{\bar{X}}, \bar{\bar{Y}}$	=	averages across measurement results obtained with *X* and *Y* measurement procedures with samples (grand averages).
$\hat{\sigma}^2_{\bar{X}}, \hat{\sigma}^2_{\bar{Y}}, \hat{\sigma}_{\bar{X}\bar{Y}}$	=	average squares and average cross-product of the deviations of the replicate averages of results of measurement obtained with the *X* and *Y* measurement procedures from the respective grand averages.
$\hat{\lambda}$	=	ratio of the variances of random errors of the two measurement procedures (within-run or repeatability when data are collected in a single run).

The constant, measurand-level–independent, random error variance estimates, $\hat{\sigma}^2(\varepsilon_X), \hat{\sigma}^2(\varepsilon_Y)$, are calculated as follows[2] (the equations are modified for the same numbers of replicates, *R*, for both measurement procedures, *X* and *Y*, and each of N specimens):

$$\hat{\sigma}^2(\varepsilon_X) = \frac{1}{\mathrm{N}(R-1)} \sum_{i=1}^{\mathrm{N}} \sum_{j=1}^{R} (X_{ij} - \bar{X}_i)^2 \tag{E15}$$

$$\hat{\sigma}^2(\varepsilon_Y) = \frac{1}{\mathrm{N}(R-1)} \sum_{i=1}^{\mathrm{N}} \sum_{k=1}^{R} (Y_{ik} - \bar{Y}_i)^2 \tag{E16}$$

Each of the above variances has N(R – 1) degrees of freedom.

The variances of the averages of *R* replicate results of measurement are *R* times smaller than the variances of the individual results given in equations (E15) and (E16):

$$\hat{\sigma}^2(\varepsilon_{\bar{X}}) = \frac{1}{\mathrm{N}R(R-1)} \sum_{i=1}^{\mathrm{N}} \sum_{j=1}^{R} (X_{ij} - \bar{X}_i)^2 \tag{E17}$$

$$\hat{\sigma}^2(\varepsilon_{\bar{Y}}) = \frac{1}{\mathrm{N}R(R-1)} \sum_{i=1}^{\mathrm{N}} \sum_{k=1}^{R} (Y_{ik} - \bar{Y}_i)^2 \tag{E18}$$

The equations for the estimates of the variances of the intercept, σ_a^2, and slope, σ_b^2, and their covariance, σ_{ab}, in Deming regression (large sample size approximation) are as follows (modified from Miller)[1]:

$$\hat{\sigma}_a^2 = \frac{1}{\mathrm{N}} \left[\hat{\sigma}^2_{\bar{Y}} - 2b\hat{\sigma}_{\bar{X}\bar{Y}} + b^2 \hat{\sigma}^2_{\bar{X}} + \frac{\bar{\bar{X}}^2 b^2}{\sigma^2_{\bar{X}\bar{Y}}} \left(\hat{\sigma}^2_{\bar{X}} \hat{\sigma}^2_{\bar{Y}} - \hat{\sigma}^2_{\bar{X}\bar{Y}} \right) \right] \tag{E19}$$

$$\hat{\sigma}_b^2 = \frac{b^2}{\mathrm{N}\hat{\sigma}^2_{\bar{X}\bar{Y}}} \left(\hat{\sigma}^2_{\bar{X}} \hat{\sigma}^2_{\bar{Y}} - \hat{\sigma}^2_{\bar{X}\bar{Y}} \right) \tag{E20}$$

$$\hat{\sigma}_{ab} = -\frac{\bar{\bar{X}} b^2}{\mathrm{N}\hat{\sigma}^2_{\bar{X}\bar{Y}}} \left(\hat{\sigma}^2_{\bar{X}} \hat{\sigma}^2_{\bar{Y}} - \hat{\sigma}^2_{\bar{X}\bar{Y}} \right) \tag{E21}$$

Appendix E. (Continued)

Assuming the Y_is follow a normal (gaussian) distribution, the above variances have $N-2$ degrees of freedom, and the $100(1-\gamma)\%$ confidence intervals for the slope and intercept are:

$$a \pm t(N-2,1-\gamma/2)\hat{\sigma}_a \quad \text{(E22)}$$

$$b \pm t(N-2,1-\gamma/2)\hat{\sigma}_b \quad \text{(E23)}$$

where

σ_a, σ_b = SDs of the intercept and slope estimates found as square roots of the respective variances in equations (E19) and (E20).

$t(N-2,1-\gamma/2) = 100(1-\gamma)$ percentile of the t-distribution with $N-2$ degrees of freedom.

The estimates of the intercept and slope are correlated. Using the variances and the covariance of the estimates allows for obtaining the joint elliptical confidence region for these parameters. Description of the method of obtaining the joint confidence region is beyond the scope of this document.

The predicted bias (B_c) at a given medical decision level X_c is:

$$\hat{B}_c = a + (b-1)X_c \quad \text{(E24)}$$

The standard error for the bias can be calculated from the variances of intercept and slope, and their covariance (equations E19, E20, E21) as follows:

$$\hat{\sigma}_{\text{Bias}} = \sqrt{\hat{\sigma}_a^2 + X_c^2\hat{\sigma}_b^2 + 2X_c\hat{\sigma}_{ab}} \quad \text{(E25)}$$

The use of the above formulas for calculating $\hat{\sigma}_a$ and $\hat{\sigma}_b$ are not appropriate when the large sample approximation and other conditions mentioned in Miller are not satisfied.[1] The jackknife approach provided in Appendix H can be implemented under less restrictive conditions and is recommended in general situations. The bootstrap, repeatedly collecting N samples with replacement from the original samples, also provides a similarly less restrictive methodology to compute the standard errors.[3]

References for Appendix E

1 Miller Jr. RG. *Beyond ANOVA, basics of applied statistics*. New York, NY: Wiley; 1986:220-230.

2 Kendall M, Stuart A. *The Advanced Theory of Statistics, Volume 2: Inference and Relationship*. 4th ed. London, England: Griffin; 1979:406-407.

3 Davison AC, Hinkley DV. *Bootstrap Methods and their Application.* Cambridge, UK: Cambridge University Press; 1997.

Appendix F. Constant CV (Weighted) Deming Regression

For measurement procedures with extremely wide measuring intervals, the analytical SD is seldom constant. Rather, a proportional relationship may apply. In a situation in which proportional analytical errors for the measurement procedures are compared, the optimal approach is a weighted form of Deming regression analysis that takes into account the relationship between random error and measurand concentration. For a given sample measured by two analytical measurement procedures, *X* and *Y*:

1. $x_i = X_{\text{Target}_i} + \varepsilon_{X_i}$ (F1)

2. $y_i = Y_{\text{Target}_i} + \varepsilon_{Y_i}$ (F2)

3. $Y_{\text{Target}_i} = a + \beta X_{\text{Target}_i}$ (F3)

x_i and y_i are the measured values, and X_{Target_i} and Y_{Target_i} are the corresponding target values. ε_{X_i} and ε_{Y_i} are the random analytical error terms of the measurement procedures *X* and *Y*, α is the regression intercept, and β is the regression slope. The analytical SDs are assumed to be proportional to the target values (CV = coefficient of variation):

$\sigma_X = \text{CV}_X X_{\text{Target}}$ and $\sigma_Y = \text{CV}_Y Y_{\text{Target}}$ (F4)

Given a proportional relationship for the random errors, a weighted procedure assigns larger weights to measurements in the low range; the low-range measurements are more precise than measurements at higher concentrations that are subject to larger random errors. More specifically, distances from (x_i, y_i) to the line are inversely weighted according to the squared analytical SDs (variances) at a given concentration that express the random error. The regression line is then estimated so that the sum of squared weighted differences is minimized. The regression procedure is most conveniently performed using dedicated software. The principle of the computations is outlined below. Weighted averages, weighted sums of squares, and a weighted crossproduct are computed:

$$\bar{X}_w = \frac{\sum_{i=1}^{N} w_i\, x_i}{\sum_{i=1}^{N} w_i} \qquad \bar{Y}_w = \frac{\sum_{i=1}^{N} w_i\, y_i}{\sum_{i=1}^{N} w_i} \quad \text{(F5, F6)}$$

$$u_w = \sum_{i=1}^{N} w_i \left(x_i - \bar{X}_w\right)^2 \qquad q_w = \sum_{i=1}^{N} w_i \left(y_i - \bar{Y}_w\right)^2 \qquad p_w = \sum_{i=1}^{N} w_i \left(x_i - \bar{X}_w\right)\left(y_i - \bar{Y}_w\right) \quad \text{(F7, F8, F9)}$$

The slope and intercept are estimated as[1,2]:

$$b = \frac{\left(\lambda q_w - u_w\right) + \sqrt{\left(u_w - \lambda q_w\right)^2 + 4\lambda p^2{}_w}}{2\lambda p_w} \quad \text{(F10)}$$

$$a_0 = \bar{Y}_w - b\bar{X}_w \quad \text{(F11)}$$

Appendix F. (Continued)

Assuming a proportional relationship, the weights are obtained by an iterative approach as described.[1,2]

$$w_i = \frac{1}{\left(\frac{\hat{X}_{\text{Target}_i} + \lambda \hat{Y}_{\text{Target}_i}}{1+\lambda} \right)^2} \quad \text{(F12)}$$

It is here presumed that the ratio λ between the squared SDs (variances) for the random error components is constant throughout the measuring interval.

$$\lambda = \sigma_x^2 \Big/ \sigma_y^2 = \text{var}(\varepsilon_X) \Big/ \text{var}(\varepsilon_Y) \quad \text{(F13)}$$

λ can be based on the analytical CVs obtained from QC results, for example. Otherwise, λ can, as default, be assigned the value 1. Without any knowledge of the ratio, for some purposes, it may be desired to vary the ratio to assess the sensitivity of the Deming regression to its value.

Bias at medical decision level(s) is calculated based on the estimates of slope and intercept. The jackknife approach provided in Appendix H can be used to calculate standard errors (SEs) of regression parameters and SE of bias.

References for Appendix F

1 Linnet K. Estimation of the linear relationship between the measurements of two methods with proportional errors. *Stat Med*. 1990;9(12):1463-1473.

2 Linnet K. Evaluation of regression procedures for methods comparison studies. *Clin Chem*. 1993;39(3):424-432.

Appendix G. Passing-Bablok Regression

The method comparison procedure of Passing and Bablok allows one to describe the linear relationship ($Y = \alpha + \beta X$) between two quantitative measurement procedures. No assumption of normal (gaussian) distribution is required. The slope of the Passing-Bablok regression line is the adjusted median of all possible slopes of the lines connecting data point pairs. The intercept is the median of intercepts that can be subsequently calculated from the data point pairs using the Passing-Bablok slope just described.

Assumptions

Two measuring systems yielding quantitative values are to be compared. A total of N individual specimens will be measured. Let *i* represent one of the N specimens.

The random variables *X* and *Y* represent outcomes of two measurement systems. The random variables can be expressed as the sums of the expected values of their respective distributions and associated measurement errors.

$$x_i = x_i^* + \xi_i \quad \text{(G1)}$$

$$y_i = y_i^* + \eta_i \quad \text{(G2)}$$

where:

x_i^*, y_i^* are the expected values of *X* and *Y*, respectively, and
ξ_i, η_i are realizations of random error terms.

The following structural relationship can be modeled:

$$y_i^* = \alpha + \beta x_i^*. \quad \text{(G3)}$$

Estimating Slope and Intercept

Given a set of N sets of N ordered pairs of measurements (x_i, y_i), where $i = 1, \ldots, N$, it is possible to determine the slopes of N. Choose two lines connecting pairs of points (x_i, y_i) and (x_j, y_j), where $1 \le i < j \le N$. The slope of the line between any two such points is $S_{ij} = (y_i - y_j) / (x_i - x_j)$. Although it is possible to obtain values of *S* equal to zero or undefined *S*, the probability of doing so is very small.

The median of the set of possible slopes is a biased estimator of β. Passing and Bablok propose adjusting the median by *K*, where *K* is the number of values of S_{ij} where $S_{ij} < -1$. Ranking the slopes in ascending order ($S_{(\text{rank order})}$) the unbiased estimator of β (b) is given by:

$$b = S_{\left(\frac{N+1}{2}+K\right)} \text{ if N is odd and } b = \frac{1}{2}\left(S_{\left(\frac{N}{2}+K\right)} + S_{\left(\frac{N}{2}+1+K\right)}\right) \text{ if N is even.} \quad \text{(G4)}$$

The intercept is defined as:

$$a = \text{median}\{y_i - bx_i\}. \quad \text{(G5)}$$

Appendix G. (Continued)

Confidence Bounds

To calculate confidence bounds for β at the γ level, let $w_{\gamma/2}$ be the $(1-\gamma/2)$ quantile of the standard normal distribution. And let

$$C_\gamma = w_{\frac{\gamma}{2}} \sqrt{\frac{N(N-1)(2N-5)}{18}} \quad \text{(G6)}$$

and

$$m_1 = \frac{N - C_\gamma}{2} \text{ and } m_2 = N - m_1 + 1, \text{ where } m_1 \text{ is rounded to the nearest integer.} \quad \text{(G7, G8)}$$

Then, a confidence interval at the γ level for β is:

$$S_{(m_1+K)} \le \beta \le S_{(m_2+K)}. \quad \text{(G9)}$$

To create confidence bounds for the intercept, let b_L and b_U represent the lower and upper confidence limits for β, and a_L and a_U represent the lower and upper confidence limits for α, then:

$$a_L = \text{median}\{y_i - b_U x_i\} \text{ and} \quad \text{(G10)}$$
$$a_U = \text{median}\{y_i - b_L x_i\}. \quad \text{(G11)}$$

Discussion

Passing and Bablok note that since the values of S_{ij} are not independent, the median of the S_{ij}s will be a biased estimator of β.[1] The authors attempt to address this bias by using the adjustment K as described above. In a follow-up article, the authors discuss the performance of their procedure compared to several others and sample size needs.[2] In a further article, Bablok et al.[3] describe modifications of the original procedure discussed above, resulting in a procedure that 1) is invariant to scale changes, 2) provides for instances where the slope between points is either 0 or 1, and 3) provides an unbiased estimator of β. For further discussion of the performance of the method, see the original papers by Passing and Bablok[1]; Bablok et al.[3]; and Linnet.[4]

References for Appendix G

1 Passing H, Bablok W. A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, Part I. *J Clin Chem Clin Biochem*. 1983;21(11):709-720.

2 Passing H, Bablok W. Comparison of several regression procedures for method comparison studies and determination of sample sizes. Application of linear regression procedures for method comparison studies in clinical chemistry, Part II. *J Clin Chem Clin Biochem*. 1984;22(6):431-445.

Appendix G. (Continued)

[3] Bablok W, Passing H, Bender R, Schneider B. A general regression procedure for method transformation. Application of linear regression procedures for method comparison studies in clinical chemistry, Part III. *J Clin Chem Clin Biochem*. 1988;26(11):783-790.

[4] Linnet K. Evaluation of regression procedures for methods comparison studies. *Clin Chem*. 1993;39(3):424-432.

Appendix H. Jackknife Approach for Estimating Standard Errors for Bias and Regression Parameters

Candidate and comparative measurement procedures are compared on the same N samples. This will be referred to as the full set. Calculate Deming slope (β), intercept (α), and bias at medical decision points (B_c) as described in Appendixes E and F. The jackknife technique should not be used for Passing-Bablok regression as described in Appendix G.

Create subset 1 by removing the first sample only. This subset will have $N-1$ samples. Create subset 2 by removing the second sample only, and continue this process for each *i*-th sample until N subsets with $N-1$ samples are created.

For each subset, calculate slope (β_i), intercept (α_i), and bias at medical decision points (B_{ci}) using the formulas previously described.

For each of the parameters above, calculate the deviations of the subsets from the full set and the average deviation as:

Deviations:

$$\delta_{bi} = Nb - (N-1)b_i \quad \text{(H1)}$$

$$\delta_{ai} = Na - (N-1)a_i \quad \text{(H2)}$$

$$\delta_{Bi} = NB_c - (N-1)B_{ci} \quad \text{(H3)}$$

Average deviation:

$$\bar{\delta}_b = \frac{\sum_{i=1}^{N} \delta_{bi}}{N} \quad \text{(H4)}$$

$$\bar{\delta}_a = \frac{\sum_{i=1}^{N} \delta_{ai}}{N} \quad \text{(H5)}$$

$$\bar{\delta}_{Bc} = \frac{\sum_{i=1}^{N} \delta_{Bci}}{N} \quad \text{(H6)}$$

Calculate standard errors (SEs) as:

SE of slope:

$$\hat{\sigma}_b = \sqrt{\frac{\sum_{i=1}^{N} (\delta_{bi} - \bar{\delta}_b)^2}{N(N-1)}} \quad \text{(H7)}$$

SE of intercept:

$$\hat{\sigma}_a = \sqrt{\frac{\sum_{i=1}^{N} (\delta_{ai} - \bar{\delta}_a)^2}{N(N-1)}} \quad \text{(H8)}$$

SE of medical decision level:

$$\hat{\sigma}_{Bias} = \sqrt{\frac{\sum_{i=1}^{N} (\delta_{Bci} - \bar{\delta}_{Bc})^2}{N(N-1)}} \quad \text{(H9)}$$

Appendix H. (Continued)

For each of these estimates (slope, intercept, bias) the confidence interval of the estimate is given by:

$$CI\,(Estimate) = Estimate \pm t\,(\mathrm{N}-2,\alpha)\hat{\sigma}_{Estimate} \qquad \text{(H10)}$$

where *Estimate* is slope (*a*), intercept (*b*), or bias (*Bc*); N is the sample size; α is the significance level (typically 95%); and *t* is the Student *t* distribution critical value.

Appendix I. A Practical Example Illustrating Bias Estimation and Measurement Procedure Comparison Techniques

A manufacturer decides to perform a comparison to determine if one lot of reagent is equivalent to another lot for a measurement procedure whose measuring interval is from near zero to 100 µg/L, whereby equivalence is concluded if the bias is within ± 6% or 0.06 µg/L, whichever is greater. In this example, the same measurement procedure is used in both instances so data from the two lots will be labeled Measurement Procedure X (MP X) and Measurement Procedure Y (MP Y). After accumulating the recommended 40 samples from an external site, it is discovered that all the samples are within a 0- to 10-µg/L measurement interval. It is decided to have the site accumulate additional samples in order to cover the entire measuring interval, resulting in a total of 79 samples. Two replicates are collected, using each lot for each sample, and the average of each replicate pair is computed. The difference and the percent difference between lots are computed. The average over both lot results is used as the divisor for computing percent difference, because both are known to have similar imprecision performance and neither can be seen as a reference lot. The dataset in Table I1, used previously in Figure 5 in Section 8.3.3 and Figure 19 in Section 9.2.3 of this document, has been ranked by this average concentration (average).

Table I1. Mixed Variability Example Dataset

MP X	MP Y	Average	Order	Diff	%Diff		MP X	MP Y	Average	Order	Diff	%Diff
0.004	0.001	0.003	1	−0.003	−120.0%		1.773	1.945	1.859	41	0.172	9.3%
0.001	0.007	0.004	2	0.006	150.0%		1.917	1.991	1.954	42	0.074	3.8%
0.007	0.001	0.004	3	−0.006	−150.0%		1.978	2.084	2.031	43	0.106	5.2%
0.007	0.001	0.004	4	−0.006	−150.0%		2.315	2.373	2.344	44	0.058	2.5%
0.004	0.012	0.008	5	0.008	100.0%		2.371	2.329	2.350	45	−0.042	−1.8%
0.012	0.004	0.008	6	−0.008	−100.0%		2.681	2.623	2.652	46	−0.058	−2.2%
0.004	0.013	0.009	7	0.009	105.9%		3.034	3.580	3.307	47	0.546	16.5%
0.014	0.006	0.010	8	−0.008	−80.0%		3.287	3.348	3.318	48	0.061	1.8%
0.014	0.008	0.011	9	−0.006	−54.5%		3.469	3.472	3.471	49	0.003	0.1%
0.008	0.015	0.012	10	0.007	60.9%		4.063	3.979	4.021	50	−0.084	−2.1%
0.030	0.012	0.021	11	−0.018	−85.7%		5.186	5.264	5.225	51	0.078	1.5%
0.018	0.026	0.022	12	0.008	36.4%		5.404	5.244	5.324	52	−0.160	−3.0%
0.026	0.018	0.022	13	−0.008	−36.4%		5.243	5.529	5.386	53	0.286	5.3%
0.030	0.041	0.036	14	0.011	31.0%		6.811	6.149	6.480	54	−0.662	−10.2%
0.040	0.036	0.038	15	−0.004	−10.5%		7.215	6.815	7.015	55	−0.400	−5.7%
0.037	0.050	0.044	16	0.013	29.9%		7.792	7.961	7.877	56	0.169	2.1%
0.045	0.051	0.048	17	0.006	12.5%		8.719	8.348	8.534	57	−0.371	−4.3%
0.051	0.045	0.048	18	−0.006	−12.5%		10.365	9.885	10.125	58	−0.480	−4.7%
0.150	0.142	0.146	19	−0.008	−5.5%		11.154	11.608	11.381	59	0.454	4.0%
0.173	0.179	0.176	20	0.006	3.4%		11.878	11.588	11.733	60	−0.290	−2.5%
0.194	0.230	0.212	21	0.036	17.0%		13.001	12.864	12.933	61	−0.137	−1.1%
0.224	0.220	0.222	22	−0.004	−1.8%		13.041	13.246	13.144	62	0.205	1.6%
0.244	0.264	0.254	23	0.020	7.9%		14.037	14.152	14.095	63	0.115	0.8%
0.338	0.340	0.339	24	0.002	0.6%		14.942	14.272	14.607	64	−0.670	−4.6%
0.645	0.653	0.649	25	0.008	1.2%		14.838	14.692	14.765	65	−0.146	−1.0%
0.607	0.703	0.655	26	0.096	14.7%		16.637	14.921	15.779	66	−1.716	−10.9%
0.641	0.697	0.669	27	0.056	8.4%		17.873	16.436	17.155	67	−1.437	−8.4%
0.666	0.739	0.703	28	0.073	10.4%		18.031	16.918	17.475	68	−1.113	−6.4%
0.744	0.768	0.756	29	0.024	3.2%		17.757	21.047	19.402	69	3.290	17.0%
0.766	0.861	0.814	30	0.095	11.7%		22.538	21.096	21.817	70	−1.442	−6.6%
0.884	0.863	0.874	31	−0.021	−2.4%		24.358	22.259	23.309	71	−2.099	−9.0%
0.871	0.883	0.877	32	0.012	1.4%		23.720	23.210	23.465	72	−0.510	−2.2%
0.880	0.877	0.879	33	−0.003	−0.3%		24.655	24.996	24.826	73	0.341	1.4%
0.893	0.955	0.924	34	0.062	−6.7%		26.155	26.577	26.366	74	0.422	1.6%
1.038	0.811	0.925	35	−0.227	−24.6%		43.709	41.220	42.465	75	−2.489	−5.9%
1.090	1.000	1.045	36	−0.090	−8.6%		41.801	43.464	42.633	76	1.663	3.9%
1.200	1.479	1.340	37	0.279	20.8%		62.516	75.876	69.196	77	13.360	19.3%
1.389	1.833	1.611	38	0.444	27.6%		69.923	71.797	70.860	78	1.874	2.6%
1.774	1.729	1.752	39	−0.045	−2.6%		91.235	99.802	95.519	79	8.567	9.0%
1.767	1.772	1.770	40	0.005	0.3%							

Appendix I. (Continued)

The data are plotted in a scatter plot and in constant difference and proportional difference plots. Note that the concentration distribution of samples in this example is still not evenly spaced over the measuring interval. To expand this into a typical 100-sample manufacturer's study, additional higher concentration samples should be collected.

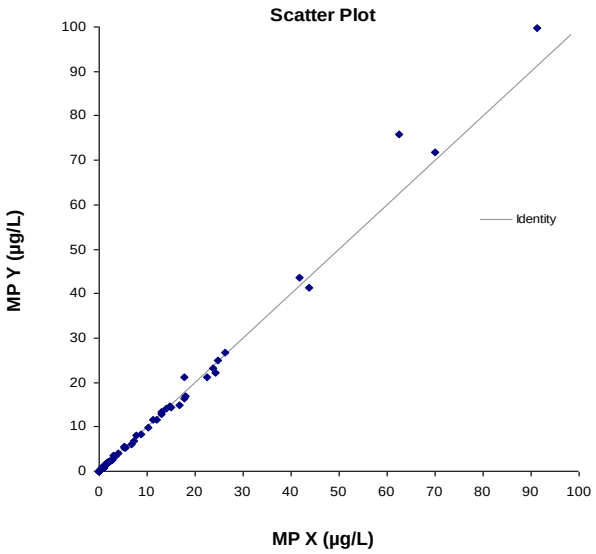

Figure I(a). Lot Comparison (see Figure 5A, Section 8.3.3
Abbreviation: MP, measurement procedure.

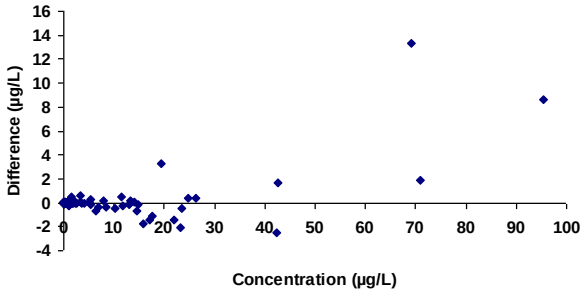

Figure I(b). Constant Difference Plot (see Figure 5B, Section 8.3.3)

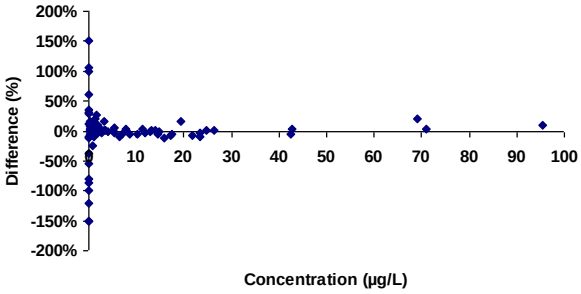

Figure I(c). Proportional Difference Plot (see Figure 5C, Section 8.3.3)

The plots visually indicate that the dataset has mixed variability with constant SD at lower concentrations and constant CV at higher concentrations. Because it is difficult to determine the concentration at which the transition between these two occurs, the difference plots are replotted with rank order as the horizontal axis.

Appendix I. (Continued)

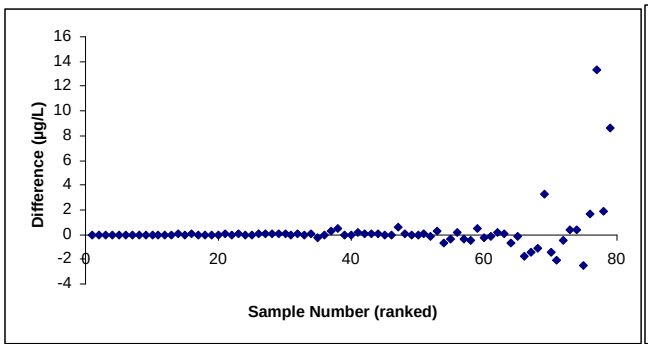

Figure I(d). Constant Difference Plot (see Figure 7A, Section 8.3.4)

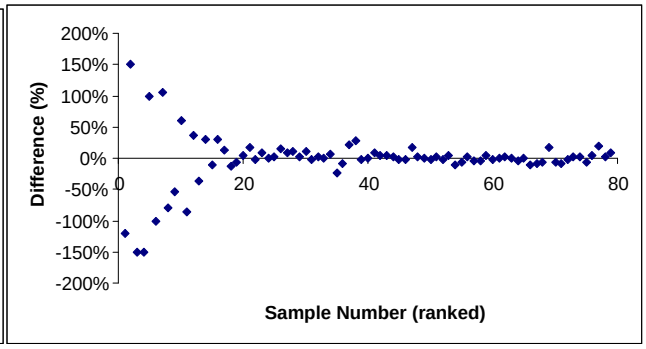

Figure I(e). Proportional Difference Plot (see Figure 7B, Section 8.3.4)

Inspecting Figure I(d), there appears to be a constant variability from sample 1 though at least sample 35. Figure I(e) displays relatively consistent proportional differences from sample 79 down to at least sample 40. Given the option of picking any sample number from 35 to 40 as the point of change of the relationship from constant to proportional, it is decided to divide the data into two equal-sized sets of results, from sample numbers 1 through 40 and from 41 to 79. The concentration at which the relationship changes from constant to proportional SD can be estimated by formal statistical analysis (called change point analysis), but is beyond the scope of this guideline.

The average difference of the low concentration dataset is 0.020 µg/L with a 95% confidence interval (CI) of −0.010–0.051 µg/L and covers the interval from 0–1.8 µg/L. The 95% CI of estimated constant bias is covered by the prespecified acceptance criterion (± 0.06 µg/L). It can therefore be concluded (with 95% confidence) that the criterion for lot equivalence was met at lower concentrations.

The average difference of the high concentration dataset is 0.43% with a 95% CI of −1.83% to 2.69% and covers the interval from 1.8–100 µg/L. The 95% CI of estimated proportional bias is covered by the prespecified acceptance criterion (± 6%). It can therefore be concluded (with 95% confidence) that the criterion for lot equivalence was met at higher concentrations.

This analysis would have been adequate for a characterization, but, for illustrative purposes, all the regression models introduced are used below to analyze the data. A review of the examples in Section 9.2 of this document demonstrates that a constant CV Deming regression would be a reasonable choice. However, given the small number of influential, high concentration samples, a Passing-Bablok regression would be the best choice. Technical discussions on the suitability of various regression techniques are referenced.[1-4]

Appendix I. (Continued)

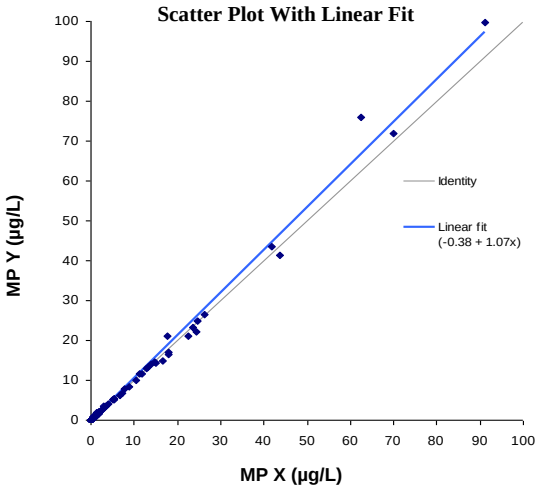

Figure I(f). Scatter Plot With Ordinary Linear Regression (OLR) Fit
Abbreviation: MP, measurement procedure.

The OLR fit demonstrates two difficulties with this technique for this dataset. First, the three highest concentration data points are so influential that the line is forced through the middle of them, regardless of the other 76 points, resulting in an estimate of positive bias (slope = 1.07). Second, the resulting pivot of line causes a low intercept, which misses most of the near-zero results (intercept = −0.36 µg/L).

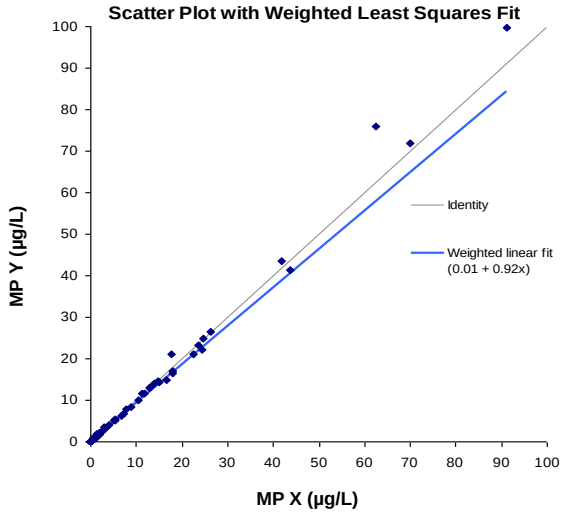

Figure I(g). Scatter Plot With Weighted Least Squares (WLS) Fit
Abbreviation: MP, measurement procedure.

The WLS regression solves the problem at the low end of the measurement interval (intercept = 0.01 µg/L), but the heightened influence of these low results causes the fitted line to miss all of the data points above 30 µg/L, resulting in a negative bias estimate (slope = 0.92).

Appendix I. (Continued)

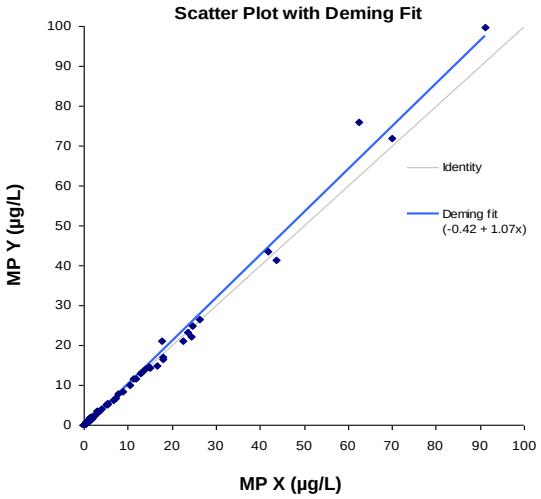

Figure I(h). Scatter Plot With Deming Fit
Abbreviation: MP, measurement procedure.

The Deming fit (nonweighted) has the same difficulties with datasets displaying proportional variability as the OLR fit. The high concentration points are very influential (slope = 1.07), which also causes a poor fit through low concentration points (intercept = −0.42 µg/L).

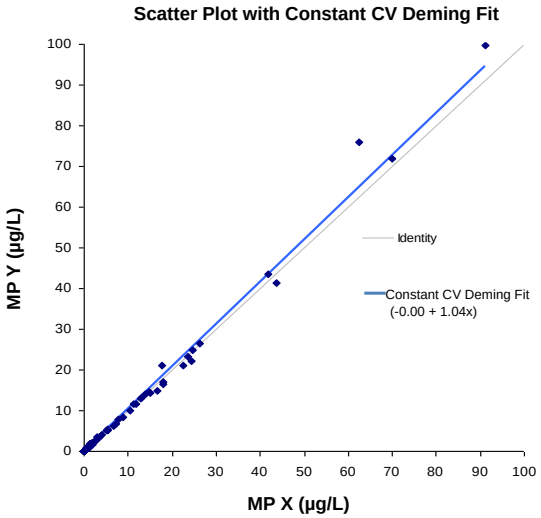

Figure I(i). Scatter Plot With Constant CV Deming Fit (see Figure 19A, Section 9.2.3)
Abbreviation: MP, measurement procedure.

The constant CV Deming fit, much like the WLS fit, ensures that the line will be drawn through the lowest concentration points (intercept = 0.00 µg/L). However, the constant CV Deming fit is not as heavily influenced by the few high concentration points (slope = 1.04).

Appendix I. (Continued)

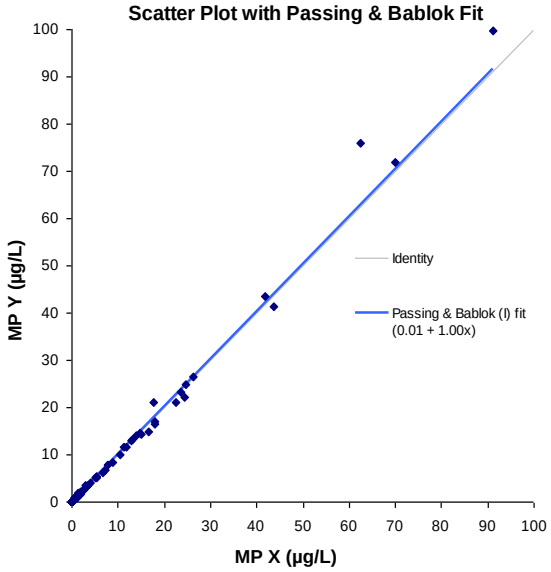

Figure I(j). Passing-Bablok Algorithm I Fit (see Figure 19B, Section 9.2.3)
Abbreviation: MP, measurement procedure.

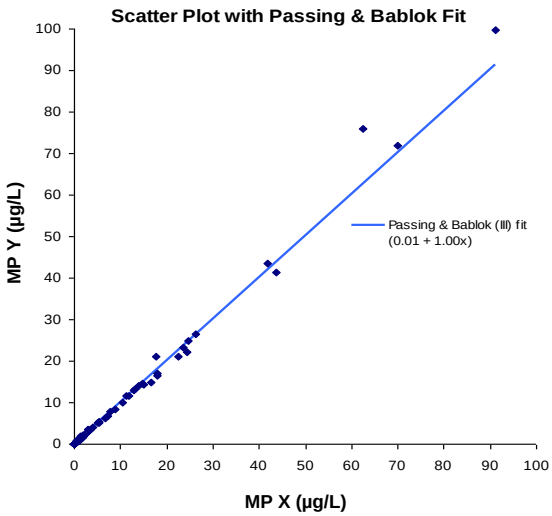

Figure I(k). Passing-Bablok Algorithm III Fit
Abbreviation: MP, measurement procedure.

The two Passing-Bablok algorithms[1,3] (see Passing and Bablok Part I and Part III references from Appendix G) provide the same estimates of slope (1.00, 95% CI, 0.98–1.02 µg/L) and intercept (0.01 µg/L, 95% CI, −0.01–0.01 µg/L). Neither is unduly influenced by either the low or the high concentration points. The slope 95% CI of 0.98 to 1.02 is equivalent to a 95% CI of proportional bias of −2% to +2%. This CI is covered by the prespecified acceptance criterion of ± 6%. It can therefore be concluded (with greater than 95% confidence) that the criterion for lot equivalence has been met.

This equation could be used to estimate the bias between the two measurement procedures at any concentration of interest. Assuming a medical decision point of 5 µg/L for Measurement Procedure X and the unrounded estimates for slope and intercept, the resultant estimate for Measurement Procedure Y would be 0.0055 + 1.0028 • 5 = 5.019 µg/L. Expressed as a proportional bias, this is (5.019 – 5.000) / average (5.000, 5.019) = 0.37%.

While the jackknife technique described in Appendix H is very amenable to providing the 95% CI of the bias estimate parametric regression methods such as OLR or Deming, using this technique for Passing-Bablok is 0.32% to 0.39%. Such an unrealistically small interval can be obtained for this nonparametric method because the exclusion of an individual point has little effect on the outcome. For Passing-Bablok, a more realistic result is obtained using the bootstrap technique, mentioned in Section 9.3 of this document. When 79 samples were selected with replacement from the sample population for 1000 individual regressions, the 95% distribution interval of bias estimates was found to be −2.02% to +1.94%. This interval is covered by the prespecified acceptance criterion of ± 6%. It can therefore be concluded (with greater than 95% confidence) that the criterion for lot equivalence has been met at this medical decision point.

References for Appendix I

1 Passing H, Bablok W. A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, Part I. *J Clin Chem Clin Biochem*. 1983;21(11):709-720.

Appendix I. (Continued)

[2] Passing H, Bablok W. Comparison of several regression procedures for method comparison studies and determination of sample sizes. Application of linear regression procedures for method comparison studies in clinical chemistry, Part II. *J Clin Chem Clin Biochem*. 1984;22(6):431-445.

[3] Bablok W, Passing H, Bender R, Schneider B. A general regression procedure for method transformation. Application of linear regression procedures for method comparison studies in clinical chemistry, Part III. *J Clin Chem Clin Biochem*. 1988;26(11):783-790.

[4] Linnet K. Evaluation of regression procedures for methods comparison studies. *Clin Chem*. 1993;39(3):424-432.

Appendix J. Example Datasets

The following tables represent data from several figures within EP09. The appropriate Appendix J table is referenced parenthetically in each figure legend.

Table J1. Constant SD Example Dataset 1

Sample	Comparative MP (µg/L)	Candidate MP (µg/L)		Sample	Comparative MP (µg/L)	Candidate MP (µg/L)
1	20.379	22.331		21	419.455	428.058
2	34.751	49.751		22	439.121	445.067
3	60.277	69.165		23	462.059	464.932
4	83.777	93.426		24	472.317	474.433
5	106.269	107.723		25	504.759	511.372
6	116.743	129.249		26	520.683	526.851
7	146.795	140.653		27	534.388	549.849
8	161.256	164.652		28	564.996	567.208
9	178.083	185.256		29	590.481	577.021
10	191.946	215.370		30	599.080	602.060
11	217.536	226.060		31	623.936	622.545
12	235.636	251.961		32	649.103	646.031
13	259.064	266.115		33	657.008	666.813
14	261.709	272.312		34	680.382	689.887
15	287.760	313.284		35	696.740	694.387
16	326.337	329.828		36	704.163	717.294
17	347.114	351.629		37	734.406	749.166
18	351.462	364.261		38	755.933	765.697
19	375.992	387.253		39	768.454	781.376
20	403.530	415.137		40	801.763	810.653

Table J2. Constant CV Example Dataset 1

Sample	Comparative MP (µg/L)	Candidate MP (µg/L)		Sample	Comparative MP (µg/L)	Candidate MP (µg/L)
1	0.881	1.027		21	108.408	120.101
2	2.872	3.405		22	148.179	148.320
3	4.975	6.125		23	153.802	156.514
4	8.351	7.582		24	159.521	189.778
5	9.753	11.110		25	198.730	174.007
6	14.450	11.774		26	203.215	203.802
7	18.552	14.278		27	215.483	184.727
8	21.520	20.553		28	227.755	250.626
9	23.481	23.890		29	233.649	293.351
10	26.710	27.532		30	298.821	314.580
11	33.259	38.645		31	276.827	291.530
12	45.907	38.764		32	362.759	378.737
13	36.180	41.887		33	439.989	430.531
14	49.853	54.300		34	434.477	402.285
15	46.330	58.702		35	572.399	444.363
16	54.798	65.969		36	689.940	461.515
17	78.002	66.507		37	538.316	628.966
18	85.464	83.946		38	612.061	763.073
19	100.030	93.344		39	734.508	712.341
20	97.877	114.151		40	785.566	670.871

Appendix J. (Continued)

Table J3. Constant CV Example Dataset 2

Sample	Comparative MP (µg/L)	Candidate MP (µg/L)		Sample	Comparative MP (µg/L)	Candidate MP (µg/L)
1	0.998	1.271		21	17.682	19.790
2	1.851	1.664		22	17.976	20.930
3	2.564	2.858		23	20.654	23.438
4	2.786	3.046		24	19.805	24.340
5	3.508	3.905		25	25.907	20.946
6	4.613	5.378		26	27.394	29.873
7	6.254	5.584		27	27.928	30.175
8	6.456	6.774		28	32.294	33.176
9	7.772	6.497		29	31.654	34.473
10	7.427	7.593		30	37.321	34.472
11	8.013	8.978		31	38.867	39.258
12	9.964	8.305		32	34.828	43.839
13	10.498	11.172		33	43.114	38.103
14	10.104	13.164		34	40.274	45.428
15	12.476	11.694		35	75.226	68.413
16	15.932	12.349		36	109.740	141.146
17	14.889	14.752		37	166.803	166.030
18	16.681	18.071		38	178.471	227.842
19	16.600	18.433		39	379.574	479.814
20	16.921	19.093		40	893.271	734.152

Table J4. Constant CV Example Dataset With Outlier

Sample	Comparative MP (µg/L)	Candidate MP (µg/L)		Sample	Comparative MP (µg/L)	Candidate MP (µg/L)
1	0.881	1.078		21	108.408	126.106
2	2.872	3.575		22	148.179	155.735
3	4.975	6.125		23	153.802	164.340
4	8.351	7.961		24	159.521	199.267
5	9.753	11.666		25	198.730	182.707
6	14.450	12.362		26	203.215	213.992
7	18.552	14.992		27	215.483	193.963
8	21.520	21.581		28	227.755	263.157
9	23.481	25.084		29	233.649	308.018
10	26.710	28.909		30	298.821	330.309
11	33.259	40.577		31	276.827	306.107
12	45.907	40.702		32	362.759	397.674
13	36.180	43.981		33	439.989	452.057
14	49.853	635.000		34	434.477	422.399
15	46.330	61.638		35	572.399	466.581
16	54.798	69.268		36	689.940	484.590
17	78.002	69.833		37	538.316	660.415
18	85.464	88.143		38	612.061	801.227
19	100.030	98.011		39	734.508	747.958
20	97.877	119.859		40	785.566	704.414

Appendix J. (Continued)

Table J5. Constant SD Example Dataset With Outlier

Sample	Comparative MP (mg/L)	Candidate MP (mg/L)		Sample	Comparative MP (mg/L)	Candidate MP (mg/L)
1	10.041	9.973		21	13.270	13.203
2	10.184	9.943		22	13.794	13.264
3	10.121	12.442		23	13.358	13.840
4	10.520	10.263		24	14.287	14.026
5	11.042	10.506		25	14.203	14.015
6	11.298	11.040		26	14.150	13.651
7	10.514	10.968		27	14.080	14.187
8	10.995	11.305		28	14.680	14.767
9	10.626	11.079		29	14.404	14.597
10	11.434	11.956		30	14.932	14.464
11	11.890	11.925		31	15.146	15.119
12	12.161	11.875		32	15.132	14.948
13	12.274	12.228		33	15.525	15.128
14	11.927	12.356		34	15.196	15.671
15	12.469	11.674		35	15.508	15.722
16	12.647	12.200		36	15.824	15.758
17	12.499	12.422		37	16.130	15.991
18	13.154	12.239		38	15.925	16.492
19	13.449	12.656		39	16.161	16.600
20	12.804	12.996		40	16.300	16.511

Table J6. Constant SD Example Dataset 2

Sample	Comparative MP (mg/L)	Candidate MP (mg/L)		Sample	Comparative MP (mg/L)	Candidate MP (mg/L)
1	10.041	10.273		21	13.270	13.503
2	10.184	11.243		22	13.794	13.564
3	10.121	9.742		23	13.358	14.140
4	10.520	10.563		24	14.287	14.326
5	11.042	10.806		25	14.203	14.315
6	11.298	11.340		26	14.150	13.951
7	10.514	11.268		27	14.080	14.487
8	10.995	11.605		28	14.680	15.067
9	10.626	11.379		29	14.404	14.897
10	11.434	12.256		30	14.932	14.764
11	11.890	12.225		31	15.146	15.419
12	12.161	12.175		32	15.132	15.248
13	12.274	12.528		33	15.525	15.428
14	11.927	12.656		34	15.196	15.971
15	12.469	11.974		35	15.508	16.022
16	12.647	12.500		36	15.824	16.058
17	12.499	12.722		37	16.130	16.291
18	13.154	12.539		38	15.925	16.792
19	13.449	12.956		39	16.161	16.900
20	12.804	13.296		40	16.300	16.811

Abbreviations: CV, coefficient of variation; MP, measurement procedure; SD, standard deviation.

This page is intentionally left blank.

The Quality Management 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 quality management system approach applies a core set of "quality system essentials" (QSEs), basic to any organization, to all operations in any health care service's path of workflow (ie, 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 QSEs are as follows:

Organization	Personnel	Process Management	Nonconforming Event Management
Customer Focus	Purchasing and Inventory	Documents and Records	Assessments
Facilities and Safety	Equipment	Information Management	Continual Improvement

EP09-A3 addresses the QSE indicated by an "X." For a description of the other documents listed in the grid, please refer to the Related CLSI Reference Materials section, beginning on the following page.

Organization	Customer Focus	Facilities and Safety	Personnel	Purchasing and Inventory	Equipment	Process Management	Documents and Records	Information Management	Nonconforming Event Management	Assessments	Continual Improvement
		I/LA21 M29	I/LA21		I/LA21	X EP28 EP31 EP05 EP06 EP07 EP12 EP14 EP15 EP17 EP21 I/LA21 I/LA28	I/LA21	I/LA21		I/LA21	I/LA21

Path of Workflow

A path of workflow is the description of the necessary processes to deliver the particular product or service that the organization or entity provides. A laboratory path of workflow consists of the sequential processes: preexamination, examination, and postexamination and their respective sequential subprocesses. All laboratories follow these processes to deliver the laboratory's services, namely quality laboratory information.

EP09-A3 does not address any of the clinical laboratory path of workflow steps. For a description of the documents listed in the grid, please refer to the Related CLSI Reference Materials section, beginning on the following page.

Preexamination				**Examination**			**Postexamination**	
Examination ordering	Sample collection	Sample transport	Sample receipt/processing	Examination	Results review and follow-up	Interpretation	Results reporting and archiving	Sample management
EP31		EP31	I/LA28	I/LA28	EP31		I/LA28	

Related CLSI Reference Materials*

EP05-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; as well as manufacturers' guidelines for establishing claims.

EP06-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.

EP07-A2 **Interference Testing in Clinical Chemistry; Approved Guideline—Second Edition (2005).** This document provides background information, guidance, and experimental procedures for investigating, identifying, and characterizing the effects of interfering substances on clinical chemistry test results.

EP12-A2 **User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline—Second Edition (2008).** This document provides a consistent approach for protocol design and data analysis when evaluating qualitative diagnostic tests. Guidance is provided for both precision and method-comparison studies.

EP14-A2 **Evaluation of Matrix Effects; Approved Guideline—Second Edition (2005).** This document provides guidance for evaluating the bias in analyte measurements that is due to the sample matrix (physiological or artificial) when two measurement procedures are compared.

EP15-A2 **User Verification of Performance for Precision and Trueness; Approved Guideline—Second Edition (2006).** 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.

EP17-A2 **Evaluation of Detection Capability for Clinical Laboratory Measurement Procedures; Approved Guideline—Second Edition (2012).** This document provides guidance for evaluation and documentation of the detection capability of clinical laboratory measurement procedures (ie, limits of blank, detection, and quantitation), for verification of manufacturers' detection capability claims, and for the proper use and interpretation of different detection capability estimates.

EP21-A **Estimation of Total Analytical Error for Clinical Laboratory Methods; Approved Guideline (2003).** This document provides manufacturers and end users with a means to estimate total analytical error for an assay. A data collection protocol and an analysis method that can be used to judge the clinical acceptability of new methods using patient specimens are included. These tools can also monitor an assay's total analytical error by using quality control samples.

EP28-A3c **Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline—Third Edition (2010).** This document contains guidelines for determining reference values and reference intervals for quantitative clinical laboratory tests. A CLSI-IFCC joint project.

EP31-A-IR **Verification of Comparability of Patient Results Within One Health Care System; Approved Guideline (Interim Revision) (2012).** This document provides guidance on how to verify comparability of quantitative laboratory results for individual patients within a health care system. A CLSI-IFCC joint project.

I/LA21-A2 **Clinical Evaluation of Immunoassays; Approved Guideline—Second Edition (2008).** This document addresses the need for clinical evaluation of new immunoassays and new applications of existing assays, as well as multiple assay formats and their uses. As a guide to designing and executing a clinical evaluation, this document will aid developers of "in-house" assays for institutional use, developers of assays used for monitoring pharmacologic effects of new drugs or biologics, and clinical and regulatory personnel responsible for commercializing products.

* CLSI documents are continually reviewed and revised through the CLSI consensus process; therefore, readers should refer to the most current editions.

Related CLSI Reference Materials (Continued)

I/LA28-A2 **Quality Assurance for Design Control and Implementation of Immunohistochemistry Assays; Approved Guideline—Second Edition (2011).** This document provides guidelines for the development of validated diagnostic, prognostic, and predictive immunohistochemical assays.

M29-A3 **Protection of Laboratory Workers From Occupationally Acquired Infections; Approved Guideline—Third Edition (2005).** Based on US regulations, this document provides guidance on the risk of transmission of infectious agents by aerosols, droplets, blood, and body substances in a laboratory setting; specific precautions for preventing the laboratory transmission of microbial infection from laboratory instruments and materials; and recommendations for the management of exposure to infectious agents.

Active Membership
(As of 1 August 2013)

Industry and Large Commercial Laboratories

Abbott Laboratories (IL)
Abbott Point of Care Inc. (NJ)
AdvaMed (DC)
Ariosa Diagnostics (CA)
ARUP Laboratories (UT)
Astellas Pharma (IL)
AstraZeneca Pharmaceuticals (MA)
Astute Medical, Inc. (CA)
Axis-Shield PoC AS (United Kingdom [GB])
Bayer Healthcare, LLC Diagnostic Division (KS)
BD (NJ)
BD Biosciences - San Jose, CA (CA)
Beckman Coulter, Inc. (PA)
Bioanalyse, Ltd. (Turkey)
Biohit Oyj. (Finland)
bioMeríeux, Inc. (MO)
Bio-Rad Laboratories, Inc. (CA)
Canon U.S. Life Sciences, Inc. (MD)
Cempra Pharmaceuticals, Inc. (NC)
Cepheid (CA)
Cerexa, Inc. (CA)
Clinical Reference Laboratory (KS)
Cubist Pharmaceuticals, Inc. (MA)
Diagnostica Stago (NJ)
DX Assays Pte Ltd. (Malaysia)
Eiken Chemical Company, Ltd. (Japan)
Elanco Animal Health (IN)
Enzo Clinical Labs (NY)
Eurofins Medinet (VA)
Exosome Diagnostics, Inc. (MN)
Greiner Bio-One GmbH (Austria)
Greiner Bio-One Inc. (NC)
Himedia Labs Ltd (India)
Hologic, Inc. (MA)
Insmed Incorporated (NJ)
Instrumentation Laboratory (MA)
Intuity Medical (CA)
Johnson & Johnson Pharmaceutical Research & Develop., L.L.C. (NJ)
Kaiser Permanente (CA)
Laboratory Corporation of America (NC)
Laboratory Specialists, Inc. (OH)
Life Laboratories (MA)
LifeLabs (Canada)
LifeLabs Medical Laboratory Services (Canada)
LipoScience, Inc. (NC)
Mbio Diagnostics, Inc. (CO)
Merck & Company, Inc. (NJ)
Microbiologics (MN)
Micromyx, LLC (MI)
Micropoint Bioscience, Inc. (CA)
Nihon Kohden Corporation (Japan)
Nissui Pharmaceutical Co., Ltd. (Japan)
Nova Biomedical Corporation (MA)
NovaBiotics (United Kingdom [GB])
Novartis Institutes for Biomedical Research (CA)
Optimer Pharmaceuticals, Inc. (CA)
Ortho-Clinical Diagnostics, Inc. (NY)
Oxyrase, Inc. (OH)
PathCare Pathology Laboratory (South Africa)
PerkinElmer Genetics, Inc. (PA)
PerkinElmer, Wallac Oy (Finland)
Pfizer Inc (PA)
Phadia AB (Sweden)
Philips Healthcare Incubator (Netherlands)
QML Pathology (Australia)
Quest Diagnostics, Incorporated (CA)
Quotient Bioresearch Ltd. (United Kingdom [GB])
Roche Diagnostics, Inc. (Spain)
Sanofi Pasteur (PA)
Sarstedt, Inc. (NC)
Sekisui Diagnostics (MA)
Seventh Sense Biosystems (MA)
Siemens Healthcare Diagnostics Inc. (IL)
Sonic Healthcare USA (TX)
SRL Limited (India)
Streck Laboratories, Inc. (NE)
Sysmex America, Inc. (IL)
Sysmex Corporation - Japan (Japan)
Tetraphase Pharmaceuticals (MA)
The Medicines Company (Canada)
Theranos (CA)
Thermo Fisher Scientific (CA)
Thermo Fisher Scientific (OH)
Thermo Scientific Microbiology Sdn Bhd (Malaysia)
Ventana Medical Systems Inc. (AZ)
Verinata Health, Inc. (CA)
Viracor-IBT Reference Laboratory (MO)
XDx, Inc. (CA)

Health Care Professions/Government

10th Medical Group (CO)
14 MDSS/SGSL (MS)
436 Medical Group (DE)
48th Medical Group (AE)
51 MDSS/ Laboratory (AP)
65th Medical Group/SGSL (AE)
673rd Medical Group (AK)
82 MDG/SGSCL Sheppard AFB (TX)
Aberdeen Royal Infirmary (United Kingdom [GB])
Academisch Ziekenhuis-VUB (Belgium)
Academy of Medical Laboratory Sciences (Ireland)
ACG (Colombia)
ACL Laboratories (IL)
ACL Laboratories (WI)
ADNOC Medical Center (United Arab Emirates)
Adventist Health System (FL)
Adventist Hinsdale Hospital (IL)
Adventist Medical Center (OR)
Affiliated Hospital of Nantong University (China)
Affiliated Laboratory, Inc. (ME)
AFRIMS (Thailand)
Aga Khan University Hospital (Pakistan)
Aggredyne, Inc. (TX)
AHS Morristown (NJ)
Akron Children's Hospital (OH)
Al Hada Armed Forces Hospital/TAIF/KSA (Saudi Arabia)
Alamance Regional Medical Center (NC)
Alameda County Medical Center (CA)
Alaska Native Medical Center (AK)
Alaska Regional Hospital (AK)
Alaska State Public Health Laboratories (AK)
Albany College of Pharmacy & Health Sciences (NY)
Albany Medical Center Hospital (NY)
Albemarle Hospital (NC)
Albert Einstein Medical Center (PA)
Alberta Health Services (Canada)
Alexandra Health Pte Ltd (Singapore)
Alfred Pathology Service (Australia)
All Children's Hospital (FL)
Alliance Community Hospital (OH)
Alliance Rehab Services (CA)
Allina Labs - 13201 (MN)
Alpena Regional Medical Center (MI)
Alta Bates Summit Medical Center (CA)
Altru Health Systems (ND)
Alvarado Hospital Medical Center Laboratory (CA)
Alverno Clinical Laboratories, Inc. (IN)
American Association for Clinical Chemistry (DC)
American Association for Laboratory Accreditation (MD)
American Hospital Dubai (United Arab Emirates)
American Medical Laboratories (Israel)
American Medical Technologists (VA)
American Society for Microbiology (DC)
American Society of Phlebotomy Technicians (SC)
American Type Culture Collection (VA)
Ameripath Indiana (IN)
Ampath (South Africa)
Analisis Clinicos ML SRL (Peru)
Analytisch Diagnostisch Centrum N.V. (Curaçao)
Ann & Robert H. Lurie Children's Hospital of Chicago (IL)
Anna Jaques Hospital (MA)
Anne Arundel Medical Center (MD)
Anson General Hospital (Canada)
Appalachian Regional Healthcare System (NC)
Arhus Universitets Hospital (Denmark)
Arizona State Health Laboratory (AZ)
Arkansas Children's Hospital (AR)
Arkansas Dept of Health (AR)
Arnot Ogden Medical Center Laboratory (NY)
Arrowhead Regional Medical Center (CA)
Asan Medical Center (Korea, Republic of)
Asante Health System (OR)
Asesores Y Consultores En Laboratorios SCU (Mexico)
Ashe Memorial Hospital (NC)
Asiri Group of Hospitals Ltd. (Sri Lanka)
Aspen Valley Hospital (CO)
ASPETAR (Qatar Orthopedic and Sports Medicine Hospital) (Qatar)
Aspirus Wausau Hospital (WI)
Associação Das Pioneiras Sociais (Brazil)
Association of Public Health Laboratories (MD)
Atlantic Diagnostics Laboratories (PA)
Atlanticare Regional Medical Center (NJ)
Audie L. Murphy VA Hospital (TX)
Augusta Health (VA)
Aultman Hospital (OH)
Austin Diagnostic Clinic (TX)
Austin Health (Australia)
Austin Regional Clinic, P.A. (TX)
Austin State Hospital (TX)
Avera Heart Hospital of South Dakota (SD)
Avera McKennan Laboratory (SD)
AZ Sint Maarten (Belgium)
AZ Sint-Jan (Belgium)
AZ Sint-Lucas Hospital (Belgium)
Azienda Ospedale Di Lecco (Italy)
B.B.A.G. Ve U. AS., Duzen Laboratories (Turkey)
Baptist Health Medical Center (FL)
Baptist Health Medical Center-Little Rock (AR)
Baptist Health System (TX)
Baptist Hospital East (KY)
Baptist Hospital Laboratory (FL)
Baptist Hospital of Miami (FL)
Baptist Memorial Health Care Corporation - Hospital Laboratories Works (TN)
Barnes-Jewish Hospital (MO)
Bassett Army Community Hospital (AK)
Bassett Healthcare (NY)
Basurto Hospital (Spain)
Baton Rouge General (LA)
Baxter Regional Medical Center (AR)
Bay Area Hospital (OR)
Bay Medical Center (FL)
BayCare Health System (FL)
Bayhealth Medical Center-Kent General Hospital (DE)
Baylor Health Care System (TX)
Bayou Pathology, APMC (LA)
Baystate Medical Center (MA)
BC Biomedical Laboratories (Canada)
BC Centre for Disease Control (Canada)
Beaufort Delta Health and Social Services Authority (Canada)
Beebe Medical Center (DE)
Bellin Hospital (WI)
Beloit Memorial Hospital (WI)
Berkshire Medical Center (MA)
Berwick Hospital Center (PA)
Beth Goldstein Consultant (PA)
Bethesda Memorial Hospital (FL)
Billings Clinic (MT)
Biodesign Institute At ASU (AZ)
Bio-Reference Laboratories (NJ)
Blanchard Valley Hospital (OH)
BloodCenter of Wisconsin (WI)
Blount Memorial Hospital (TN)
Blue Mountain Health System (PA)
Blue Ridge Regional Hospital (NC)
Boca Raton Community Hospital (FL)
Bon Secours Baltimore Health System (MD)
Bon Secours Health Partners (VA)
Bon Secours Hospital (Ireland)
Boulder Community Hospital (CO)
Bozeman Deaconess Laboratory (MT)
Braintree Rehabilitation Hospital (MA)
Brandywine Hospital (PA)
Brant Community Healthcare System/Brant General Hospital (Canada)
Brazosport Regional Health System (TX)
Breathitt Veterinary Center, Murray State University (KY)
Brian All Good Community Hospital/121 Combat (CA)
Bridgeport Hospital (CT)
Bristol Hospital (CT)
British Columbia Institute of Technology (Canada)
Brockville General Hospital (Canada)
Bronson Methodist Hospital (MI)
Broward General Medical Center (FL)
Brownwood Regional Medical Center (TX)
Bryan LGH Medical Center (NE)
BSA Health System (TX)
Buena Vista Regional Medical Center (IA)
Bumrungrad Hospital (Thailand)
C. Gregory Bowling, MD APMC (LA)
Cadham Provincial Laboratory-MB Health (Canada)
California Department of Public Health (CA)
California Pacific Medical Center (CA)
Cambridge Health Alliance (MA)
Cambridge Life Science (United Kingdom [GB])
Camden Clark Memorial Hospital (WV)
Cameron Regional Medical Center (MO)
Campbellford Memorial Hospital (Canada)
Canadian Science Center for Human and Animal Health (Canada)
Canadian Society for Medical Laboratory Science (Canada)
Canberra Hospital (Australia)
Cape Cod Hospital (MA)
Cape Fear Valley Medical Center Laboratory (NC)
Capital Coast Health (New Zealand)
Capital Health Regional Medical Center (NJ)
Capital Region Medical Center (MO)
Cardinal Hill Rehabilitation Hospital (KY)
Caritas Norwood Hospital (MA)
Carl R. Darnall Army Medical Center Department of Pathology (TX)
Carle Foundation Hospital (IL)
Carolinas Healthcare System (NC)
Caromont Regional Medical Center (NC)
Carpermor S.A. de C.V. (Mexico)
Carrington College (AZ)
Carroll Hospital Center (MD)
Carteret General Hospital (NC)
Cary Medical Center (ME)
Cass County Memorial Hospital (IA)
Castle Medical Center (HI)
Catholic Health Systems-Sisters of Charity Hospital (NY)
Catholic Medical Center (NH)
Cayuga Medical Center At Ithaca (NY)
CD Diagnostics, Inc. (PA)
CDC - Nigeria (Nigeria)

Cedars-Sinai Medical Center (CA)
Cedimat Medical Center (Dominican Republic)
Cellnetix Pathology & Laboratories (WA)
Center for Disease Detection (TX)
Centers for Disease Control and Prevention (GA)
Centers for Disease Control and Prevention - Ethiopia (Ethiopia)
Centers for Disease Control and Prevention – Tanzania (Tanzania, United Republic of)
Centers for Medicare & Medicaid Services (MD)
Centers for Medicare & Medicaid Services/CLIA Program (TX)
Centers for Medicare and Medicaid Services (GA)
CentraCare Health - Monticello (MN)
Central Baptist Hospital (KY)
Central Maine Medical Center (ME)
Central Ohio Primary Care Physicians (OH)
Central Pennsylvania Alliance Laboratory (PA)
Central Vermont Medical Center (VT)
Central Washington Hospital (WA)
Centre Hospitalier Anna-Laberge (Canada)
Centre Hospitalier Lyon SUD (France)
Centre Hospitalier Regional De Trois Riveras (Canada)
Centro Medico Imbanaco (Colombia)
CGH Medical Center (IL)
Chaleur Regional Hospital (Canada)
Chambersburg Hospital (PA)
Champlain Valley Physicians Hospital (NY)
Changhua Christian Hospital (Taiwan)
Changi General Hospital (Singapore)
Charleston Area Medical Center (WV)
Chatham - Kent Health Alliance (Canada)
Chebucto Medical Collection (Canada)
Chesapeake General Hospital (VA)
Chester County Hospital (PA)
Cheyenne Regional Medical Center (WY)
Chi Solutions, Inc. (MI)
Chia-Yi Chang Gung Memorial Hospital (Taiwan)
Children's Healthcare of Atlanta (GA)
Childrens Hosp.- Kings Daughters (VA)
Children's Hospital (AL)
Children's Hospital & Medical Center (NE)
Children's Hospital Boston (MA)
Childrens Hospital Los Angeles (CA)
Children's Hospital Medical Center (OH)
Children's Hospital of Central California (CA)
Children's Hospital of Philadelphia (PA)
Childrens Hospital of Wisconsin (WI)
Children's Hospitals and Clinics (MN)
Children's Medical Center (TX)
Chilton Memorial Hospital (NJ)
Chinese Committee for Clinical Laboratory Standards (China)
Chino Valley Medical Center (CA)
Christiana Care Health Services (DE)
Christus Spohn Hospital Beeville (TX)
Christus St. Patrick Hospital (LA)
CHU Sainte-Justine: Department of Microbiology and Immunology (Canada)
CHUM Hospital Saint-Luc (Canada)
Chungnam National University Hospital (Korea, Democratic People's Republic)
CHU-St. Justine (Canada)
CHW-St. Mary's Medical Center (CA)
Cibola General Hospital (NM)
Citizens Memorial Hospital (MO)
Citrus Memorial Hospital (FL)
City of Hope National Medical Center (CA)
City of Milwaukee Health Department (WI)
Clara Maass Medical Center (NJ)
Cleveland Clinic (OH)
Cleveland Regional Medical Center (NC)
Clifton Fine Hospital (NY)
Clinica Hospital San Fernando (Panama)
Clinical Hospital Merkur (Croatia/Hrvatska)
Clinical Labs of Hawaii (HI)
Clinique St. Luc (Belgium)
CLMA (IL)
CML HealthCare (Canada)
COLA (MD)
College of American Pathologists (IL)
College of Physicians and Surgeons of Alberta (Canada)
College of Physicians and Surgeons of Saskatchewan (Canada)
College of the North Atlantic (Canada)
College of Veterinary Medicine, Auburn University (AL)
Collingwood General & Marine Hospital (Canada)
Collom & Carney Clinic (TX)
Columbia Memorial Hospital (NY)
Columbia Memorial Hospital (OR)
Columbia St. Mary's Milwaukee (WI)
Columbus Regional Healthcare System (NC)
Commonwealth of Kentucky (KY)
Commonwealth of Virginia (DCLS) (VA)
Community College of Rhode Island-Flanagan Campus (RI)
Community Hospital (IN)
Community Hospital of the Monterey Peninsula (CA)
Community Medical Center (MT)
Community Medical Center (NJ)
Complexe Hospitalier de la Sagamie (Canada)
CompuNet Clinical Laboratories (OH)
Coney Island Hospital (NY)
Consultants Laboratory of WI LLC (WI)
Contra Costa Regional Medical Center (CA)
Conway Medical Center (SC)
Cook Children's Medical Center (TX)
Cookeville Regional Medical Center (TN)
Cooper University Hospital (NJ)
Cornwall Community Hospital (Canada)
Corvallis Clinic (OR)
Countess of Chester Hospital (United Kingdom [GB])
Counties Manukau District Health Board, Middlemore Hospital (New Zealand)
Covance CLS (IN)
Covenant Health Care (MI)
Covenant Medical Center (TX)
Crozer-Chester Medical Center (PA)
CSSS Alphonse-Desjardins (Canada)
CSSS Du Sud De Lanaudiere (Canada)
CSSS St-Jerome (Canada)
Cumberland Medical Center (TN)
Cyruss Tsurgeon (LA)
Dameron Hospital Association (CA)
Danbury Hospital (CT)
Darwin Health Library, NT Dept. of Health (Australia)
Daviess Community Hospital (IN)
DaVita Laboratory Services, Inc. (FL)
Dayton Children's Medical Center (OH)
Deaconess Hospital (WA)
Deaconess Hospital Laboratory (IN)
Dean Medical Center (WI)
Delaware Public Health Laboratory (DE)
Delnor Community Hospital (IL)
Delta Regional Medical Center (MS)
Denver Health & Hospital Authority (CO)
Dept. of VA Affairs: Regional Commissioners Program (TX)
Dermatopathology Northwest (WA)
DHHS NC State Lab of Public Health (NC)
Diagnostic Accreditation Program (Canada)
Diagnostic Center for Population & Animal Health (MI)
Diagnostic Laboratory Services, Inc. (HI)
Diagnostic Services of Manitoba (Canada)
Dialysis Clinic, Inc. Laboratory (TN)
DIATHERIX Laboratories, Inc. (AL)
Dimensions Healthcare System Prince George's Hospital Center (MD)
DMC University Laboratories (MI)
Doctors Hospital (FL)
Doctors Hospital (OH)
DoctorsManagement (TN)
Dokkyo Medical University Hospital (Japan)
Donalsonville Hospital (GA)
DPH - Newborn Screening Program (DE)
Dr. Soliman Fakeeh Hospital (Saudi Arabia)
Driscoll Children's Hospital (TX)
Drug Scan Inc. (PA)
DuBois Regional Medical Center (PA)
DUHS Clinical Laboratories (NC)
Duke University Medical Center (NC)
Dynacare Laboratory (WI)
Dynacare NW, Inc - Seattle (WA)
DynaLIFE (Canada)
E. A. Conway Medical Center (LA)
East Georgia Regional Medical Center (GA)
East Texas Medical Center - Tyler (TX)
East Texas Medical Center (ETMC) Henderson (TX)
East Texas Medical Center-Pittsburg (TX)
Eastern Gateway Community College (OH)
Eastern Health - Health Sciences Centre (Canada)
Eastern Health Pathology (Australia)
Eastern Ontario Regional Laboratory Association (EORLA) (Canada)
Easton Hospital (PA)
Edgerton Hospital & Health Services (WI)
Edmonds Community College (WA)
Edward Hospital (IL)
Eisenhower Army Medical Center (GA)
El Camino Hospital (CA)
Emerson Hospital Laboratory (MA)
EMH Regional Medical Center (OH)
Emory University Hospital (GA)
Emory University School of Medicine (GA)
Empire College (CA)
Ephrata Community Hospital (PA)
Erasmus University Medical Center (Netherlands)
Erlanger Health Systems (TN)
ESCMID (Switzerland)
Estes Park Medical Center (CO)
Ethiopian Health and Nutrition Research Institute (Ethiopia)
Evangelical Community Hospital (PA)
Evans Army Community Hospital (CO)
Evanston Hospital, NorthShore University HealthSystem (IL)
Excela Health Latrobe Hospital (PA)
Exempla - Saint Joseph Hospital (CO)
Exempla Lutheran Medical Center (CO)
Fairfax County Health Department (VA)
Farrer Park Hospital (Singapore)
Fauquier Hospital (VA)
Fayette County Memorial Hospital (OH)
FDA Ctr. for Devices/Rad. Health (CDRH) (MD)
Federal Medical Center (MN)
FHG- University of Applied Science-Tyrol (Austria)
Firelands Regional Medical Center (OH)
Fisher County Hospital (TX)
Fisher-Titus Memorial Hospital (OH)
Flagler Hospital Inc. (FL)
Flagstaff Medical Center (AZ)
Fletcher Allen Health Care (VT)
Florida Hospital Flagler (FL)
Flushing Hospital (NY)
Forrest General Hospital (MS)
Forsyth Medical Center (NC)
Fort Loudoun Medical Center (TN)
Fox Chase Cancer Center (PA)
Franklin Memorial Hospital (ME)
Fresno Community Hospital & Medical Center (CA)
Ft. Belvoir Community Hospital (VA)
Fundación Mexicana Para la Salud Capitulo Peninsular A.C (Mexico)
Gamma-Dynacare Laboratories (Canada)
Garden City Hospital (MI)
Gateway Regional Medical Center (IL)
Geary Community Hospital (KS)
Geisinger Medical Center (PA)
Genesis Healthcare System (OH)
Genesis Laboratory Management (NJ)
Genesis Medical Center (IL)
George Mason University (VA)
Ghent University Hospital (Belgium)
Glasgow Royal Infirmary (United Kingdom [GB])
Golden Valley Memorial Hospital (MO)
Golwilkar Metropolis (India)
Good Samaritan Hospital (IN)
Good Samaritan Hospital Medical Center (NY)
Good Shepherd Medical Center (TX)
Gottlieb Memorial Hospital (IL)
Grady Health System Laboratory (GA)
Grana S.A. (TX)
Grand River Hospital (Canada)
Grays Harbor Community Hospital (WA)
Great Plains Regional Med. Ctr. (NE)
Great River Medical Center (IA)
Greater Baltimore Medical Center (MD)
Greater Lowell Pediatrics (MA)
Green Cross Reference Laboratories (Korea, Republic of)
Greenbrier Valley Medical Center (WV)
Greensboro Pathology (NC)
Greenville Memorial Medical Campus (SC)
Greenwood Leflore Hospital (MS)
Grey Bruce Regional Health Center (Canada)
Gritman Medical Center (ID)
Group Health Cooperative (WA)
Group Health Cooperative - SCW (WI)
Grove City Medical Center (PA)
Guelph General Hospital (Canada)
Gulf Medical College Hospital & Research Centre (United Arab Emirates)
Gundersen Lutheran Medical Center (WI)
Gunnison Valley Hospital (CO)
Guthrie Clinic Laboratories (PA)
Gwinnett Medical Center (GA)
H. Lee Moffitt Cancer Center (FL)
Hagerstown Medical Laboratory (MD)
Halifax Regional Medical Center (NC)
Halton Healthcare Services (Canada)
Hamad Medical Corp-DLMP LAB QM (Qatar)
Hamad Medical Corporation (Qatar)
Hamilton Hospital (TX)
Hamilton Regional Laboratory Medicine Program - St. Joseph's (Canada)
Hampton Regional Medical Center (SC)
Hanover General Hospital (PA)

Harbor - UCLA Medical Center (CA)
Hardy Diagnostics (CA)
Harford Memorial Hospital (MD)
Harris Hospital (AR)
Harris Methodist HEB Hospital (TX)
Harris Methodist Hospital Southwest (TX)
Hartford Hospital (CT)
Harvard Vanguard Medical Associates (MA)
Hawaii Pathologists Laboratory (HI)
Hawaii State Hospital (HI)
Healdsburg District Hospital (CA)
Health Canada (Canada)
Health Diagnostic Laboratory, Inc. (VA)
Health Network Lab (PA)
Health Sciences North (Canada)
Health Waikato (New Zealand)
Healthscope Pathology (Australia)
Healthtronics Lab Solutions (PA)
Heartland Health (MO)
Helen Ellis Memorial Hospital (FL)
Helen Hayes Hospital (NY)
Helena Regional Medical Center (AR)
Hema-Quebec (Canada)
Hendrick Regional Laboratory (TX)
Hendricks Regional Health (IN)
Hennepin County Medical Center (MN)
Henrico Doctors' Hospital - Parham (VA)
Henry Ford Hospital (MI)
Henry M. Jackson Foundation for the Advancement of Military Medicine-MD (MD)
Henry M. Jackson Foundation-Brook Army Medical Ctr (BAMC) (TX)
Hi-Desert Medical Center (CA)
Highlands Medical Center (AL)
Highline Medical Center (WA)
Hillcrest Medical Center (OK)
Hinsdale Pathology Associates (IL)
Hoag Memorial Hospital Presbyterian (CA)
Holstebro Hospital (Denmark)
Holy Name Hospital (NJ)
Holy Redeemer Hospital & Medical Center (PA)
Holy Spirit Hospital (PA)
Holzer Health System (OH)
Hong Kong Accreditation Service Innovation and Technology Commission (Hong Kong)
Hong Kong Sanatorium & Hospital (Hong Kong)
Hôpital Cite de La Sante De Laval (Canada)
Hopital de Granby-CSSS Haute-Yamaska (Canada)
Hopital du Haut-Richelieu (Canada)
Hopital Maisonneuve-Rosemont (Canada)
Hopital Santa Cabrini Ospedale (Canada)
Hopital Ste - Croix, CSSS Drummond (Canada)
Hopkins County Memorial Hospital (TX)
Horizon Health Network (Canada)
Hospital Albert Einstein (Brazil)
Hospital Italiano Laboratorio Central (Argentina)
Hospital Sacre-Coeur de Montreal (Canada)
Hotel Dieu Grace Hospital Library (Canada)
Houston Medical Center (GA)
Hunt Regional Healthcare (TX)
Hunterdon Medical Center (NJ)
Huntington Memorial Hospital (CA)
Hutchinson Clinic, P.A. (KS)
Hutt Valley Health District Health Board (New Zealand)
IDEXX Reference Laboratories (Canada)
Incyte Pathology (WA)
Indiana University - Chlamydia Laboratory (IN)
Indiana University Health Bloomington Hospital (IN)
Indiana University Health Care - Pathology Laboratory (IN)
INEI-ANLIS "Dr. C. G. Malbrán" (Argentina)
Ingalls Hospital (IL)
Inova Central Laboratory (VA)
Institut Für Klinische Chemie Und Laboratoriumsmedizin Universitätsklinikum (Germany)
Institut National de Santa Publique Du Quebec Centre de Doc. - INSPQ (Canada)
Institute Health Laboratories (PR)
Institute of Laboratory Medicine Landspitali Univ. Hospital (Iceland)
Institute of Public Health (Slovenia)
Institute of Tropical Medicine Dept. of Clinical Sciences (Belgium)
Institute of Veterinary Bacteriology (Switzerland)
Instituto Nacional de Ciencias Médicas y Nutrición (Mexico)
Integrated BioBank (Luxembourg)
Integrated Diagnositcs (WA)
Integrated Regional Laboratories (HCA) (FL)
Interim LSU Hospital/Med. Center of La (LA)
Interior Health (Canada)
International Accreditation New Zealand (New Zealand)
International Health Management Associates, Inc. (IL)
Irwin Army Community Hospital (KS)
Istituto Cantonale Di Microbiologia (Switzerland)
Jack Hughston Memorial Hospital (AL)
Jackson County Memorial Hospital (OK)
Jackson Health System (FL)
Jackson Hospital & Clinic, Inc. (AL)
Jackson Purchase Medical Center (KY)
Jam Yperman Hospital (Belgium)
Jameson Memorial Hospital (PA)
Japan Assn. of Clinical Reagents Industries (Japan)
Jefferson Memorial Hospital (WV)
Jefferson Regional Medical Center (PA)
Jennings American Legion Hospital (LA)
Jersey Shore University Medical Center (NJ)
Jessa Ziekenhuis VZW (Belgium)
Jiao Tong University School of Medicine - Shanghai No. 3 People's Hospital (China)
John C. Lincoln Hospital - N.MT. (AZ)
John D. Archbold Hospital (GA)
John F. Kennedy Medical Center (NJ)
John H. Stroger, Jr. Hospital of Cook County (IL)
Johns Hopkins APL (MD)
John Muir Health (CA)
Johns Hopkins Medical Institutions (MD)
Johnson City Medical Center Hospital (TN)
Johnston Memorial Hospital (NC)
Jonathan M. Wainwright Memorial Veterans Affairs Medical Center (WA)
Jones Memorial Hospital (NY)
Jordan Valley Community Health Center (MO)
JPS Health Network (TX)
Jupiter Medical Center (FL)
Kaiser Medical Laboratory (HI)
Kaiser Permanente (GA)
Kaiser Permanente (MD)
Kaiser Permanente Colorado (CO)
Kaiser Permanente Medical Care (CA)
Kaiser Permanente San Francisco (CA)
Kaiser TPMG Medical Center (CA)
Kaleida Health Center for Laboratory Medicine (NY)
Kalispell Regional Medical Center (MT)
Kane Community Hospital (PA)
Kansas Department of Health & Environment (KS)
Kansas State University (KS)
Kantonsspital Aarau AG (Switzerland)
Kaohsiun Chang Gung Memorial Hospital (Taiwan)
Karmanos Cancer Institute (MI)
KCHL St. Elisabeth Hospital (Netherlands)
Keck Hospital of USC (CA)
Keck School of Medicine-USC (CA)
Keelung Chang Gung Memorial Hospital (Taiwan)
Keller Army Community Hospital (NY)
Kennedy Health System (NJ)
Kenora-Rainy River Reg. Lab. Program (Canada)
Kent County Memorial Hospital (RI)
Kettering Medical Center (OH)
Kindred Healthcare (KY)
King Abdulaziz Hospital (Saudi Arabia)
King Fahad Medical City (Saudi Arabia)
King Fahad Specialist Hospital-Dammam, K.S.A. (Saudi Arabia)
King Faisal Specialist Hospital & Research Center (Saudi Arabia)
King Hussein Cancer Center (Jordan)
Kingsbrook Jewish Medical Center (NY)
Kingston General Hospital (Canada)
KK Women's & Children's Hospital (Singapore)
Kuakini Health System (HI)
Kyoto University Hospital (Japan)
Lab Express (AZ)
Lab Médico Santa Luzia LTDA (Brazil)
Labor Stein + Kollegen (Germany)
Laboratorio Bueso Arias (Honduras)
Laboratorio Clinico Amadita P. de Gonzales S.A. (FL)
Laboratorio de Referencia (FL)
Laboratorio Médico De Referencia (Colombia)
Laboratory Alliance of Central New York (NY)
Laboratory Corporation of America (NJ)
Laboratory for Medical Microbiology and Infectious Diseases (Netherlands)
Laboratory Medicin Dalarna (Sweden)
Laboratory of Clinical Biology Ziekenhuis Oost-Limburg (ZOL) (Belgium)
Laboratory of Veterinary Medicine (Luxembourg)
LabPlus Auckland District Health Board (New Zealand)
LAC/USC Medical Center (CA)
Lafayette General Medical Center (LA)
Lahey Clinic (MA)
Lake Charles Memorial Hospital (LA)
Lake Health (OH)
Lake Wales Medical Center (FL)
Lakeland Regional Laboratories (MI)
Lakeland Regional Medical Center (FL)
Lakeridge Health Corporation - Oshawa Site (Canada)
Lakeview Medical Center (WI)
Lakeway Regional Medical Center (TX)
Lamb Healthcare Center (TX)
Lancaster General Hospital (PA)
Landstuhl Regional Medical Center (AE)
Lane Regional Medical Center (LA)
Langley Air Force Base (VA)
Lawrence and Memorial Hospitals (CT)
LeBonheur Children's Hospital (TN)
Leesburg Regional Medical Center (FL)
Legacy Laboratory Services (OR)
Leiden University Medical Center (Netherlands)
Lexington Medical Center (SC)
L'Hotel-Dieu de Quebec (Canada)
Licking Memorial Hospital (OH)
LifeBridge Health Sinai Hospital (MD)
LifeCare Medical Center (MN)
Little Company of Mary Hospital (IL)
Littleton Hospital (NH)
Lodi Memorial Hospital (CA)
Lompoc Valley Medical Center (CA)
London Health Sciences Center (Canada)
Long Beach Memorial Medical Center-LBMMC (CA)
Long Island Jewish Medical Center (NY)
Longmont United Hospital (CO)
Longview Regional Medical Center (TX)
Louisiana Office of Public Health Laboratory (LA)
Louisiana State University Medical Ctr. (LA)
Lourdes Health System (NJ)
Lower Bucks Hospital (PA)
Lower Mainland Laboratories (WA)
Loyola University Medical Center (IL)
Luminex Corporation (TX)
Lummi Tribal Health Center (WA)
Lutheran Hospital of Indiana Inc. (IN)
Lynchburg General (VA)
Lyndon B. Johnson General Hospital (TX)
Lyster Army Health Clinic (AL)
MA Dept. of Public Health Laboratories (MA)
Mackenzie Health (Canada)
Madigan Army Medical Center (WA)
Mafraq Hospital (United Arab Emirates)
Magee Womens Hospital of UPMC (PA)
Magnolia Regional Health Center (MS)
Main Line Clinical Laboratories, Inc. Lankenau Hospital (PA)
Maine General Medical Center (ME)
Mammoth Hospital Laboratory (CA)
Manatee Hospitals and Health (FL)
Maria Parham Medical Center (NC)
Marietta Memorial Hospital (OH)
Marin General Hospital (CA)
Marion County Public Health Department (IN)
Marquette General Hospital (MI)
Marshfield Clinic (WI)
Martha Jefferson Hospital (VA)
Martin Luther King, Jr./Drew Medical Center (CA)
Martin Memorial Health Systems (FL)
Mary Greeley Medical Center (IA)
Mary Hitchcock Memorial Hospital (NH)
Mary Washington Hospital (VA)
Massachusetts General Hospital (MA)
Massasoit Community College (MA)
Mater Health Services - Pathology (Australia)
Mayo Clinic (MN)
Mayo Clinic Health Systems in Waycross (GA)
Mayo Clinic Scottsdale (AZ)
McAlester Regional Health Center (OK)
McCullough-Hyde Memorial Hospital (OH)
McCune-Brooks Hospital (MO)
MCG Health (GA)
McKenzie-Willamette Medical Center (OR)
McLaren Northern Michigan (MI)
MCN Healthcare (CO)
Meadows Regional Medical Center (GA)
Meadville Medical Center (PA)
Med. Laboratories Duesseldorf (Germany)
Media Lab, Inc. (GA)
Medibus (Canada)
Medical Center Enterprise (AL)
Medical Center Hospital (TX)
Medical Center of Central Georgia (GA)
Medical Centre Ljubljana (Slovenia)
Medical College of Virginia Hospital (VA)
Medical Laboratories of Windsor, LTD (Canada)
Medical Laboratory Sciences Council of Nigeria (Nigeria)
Medical University Hospital Authority (SC)
Medlab Central (New Zealand)
Memorial Health System (CO)

Memorial Health Systems of East Texas (TX)
Memorial Hermann Healthcare System (TX)
Memorial Hospital at Gulfport (MS)
Memorial Hospital of Carbondale (IL)
Memorial Hospital of Rhode Island (RI)
Memorial Hospital of Union City (OH)
Memorial Medical Center (IL)
Memorial Medical Center (PA)
Memorial Medical Center (TX)
Memorial Regional Hospital (FL)
Memorial Sloan Kettering Cancer Center (NY)
Mercy Franciscan Mt. Airy (OH)
Mercy Health Center (OK)
Mercy Hospital (IA)
Mercy Hospital (MN)
Mercy Hospital Jefferson (MO)
Mercy Hospital of Tiffin (OH)
Mercy Integrated Laboratories/Mercy St. Vincent (OH)
Mercy Medical Center (CA)
Mercy Medical Center (IA)
Mercy Medical Center (MD)
Mercy Medical Center (OH)
Mercy Regional Medical Center (OH)
Methodist Dallas Medical Center (TX)
Methodist Healthcare (TN)
Methodist Hospital (TX)
Methodist Hospital of Southern California (CA)
Methodist Hospital Pathology (NE)
Methodist Medical Center (TN)
Methodist Sugarland Hospital (TX)
MetroHealth Medical Center (OH)
Metropolitan Hospital Center (NY)
Metropolitan Medical Laboratory, PLC (IA)
Miami Children's Hospital (FL)
Michigan Dept. of Community Health (MI)
Michigan State University (MI)
Microbial Research, Inc. (CO)
Microbiology Specialists, Inc. (TX)
Mid America Clinical Laboratories (IN)
Mid Michigan Medical Center - Midland (MI)
Middelheim General Hospital (Belgium)
Middlesex Hospital (CT)
Midland Memorial Hospital (TX)
Midwestern Regional Medical Center (IL)
Mile Bluff Medical Center/Hess Memorial Hospital (WI)
Milford Regional Hospital (MA)
Ministry of Health - Zambia (Zambia)
Ministry of Health and Social Welfare - Tanzania (Tanzania)
Minneapolis Community and Technical College (MN)
Minneapolis Medical Research Foundation (MN)
Minnesota Department of Health (MN)
MiraVista Diagnostics (IN)
Mission Hospitals Laboratory (NC)
Mississippi Baptist Medical Center (MS)
Mississippi Public Health Lab (MS)
Missouri State Public Health Laboratory (MO)
Mobile Infirmary Association (AL)
Modesto Memorial Hospital (CA)
MolecularMD Corp. (OR)
Monadnock Community Hospital (NH)
Mongolian Agency for Standardization and Metrology (Mongolia)
Monongahela Valley Hospital (PA)
Monongalia General Hospital (WV)
Montana Department of Public Health and Human Services (MT)
Montefiroe Medical Center (NY)
Montgomery Hospital (PA)
Montgomery Regional Hospital (VA)
Morehead Memorial Hospital (NC)
Morristown Hamblen Hospital (TN)
Mount Nittany Medical Center (PA)
Mt. Auburn Hospital (MA)
Mt. Sinai Hospital (Canada)
Mt. Sinai Hospital - New York (NY)
Mt. Sinai Hospital Medical Center (IL)
Muleshoe Area Medical Center (TX)
MultiCare Health Systems (WA)
Muskoka Algonquin Healthcare (Canada)
Naas General Hospital-NGH (Ireland)
Nacogdoches Memorial Hospital (TX)
Nanticoke Memorial Hospital (DE)
Nash General Hospital/Laboratory (NC)
Nassau County Medical Center (NY)
National Cancer Institute, CDP, NIH (MD)
National Food Institute Technical University of Denmark (Denmark)
National Health Laboratory Service C/O F&M Import & Export Services (South Africa)
National Heart Institute (Institut Jantung Negra) (Malaysia)
National Institute of Health-Maputo, Mozambique (Mozambique)
National Institutes of Health, Clinical Center (MD)
National Jewish Health (CO)
National Pathology Accreditation Advisory Council (Australia)
National Society for Histotechnology, Inc. (MD)
National University Hospital (Singapore) Pte Ltd (Singapore)
National University of Ireland, Galway (NUIG) (Ireland)
National Veterinary Institute (Sweden)
Nationwide Children's Hospital (OH)
Naval Health Clinic Charleston (SC)
Naval Hospital Lemoore (CA)
Naval Hospital Oak Harbor (WA)
Naval Medical Center Portsmouth (VA)
Naval Medical Center San Diego (CA)
NB Department of Health (Canada)
Nellis Air Force Base (NV)
Netlab SA (Ecuador)
New Brunswick Community College (Canada)
New Brunswick Provincial Veterinary Laboratory (Canada)
New England Baptist Hospital (MA)
New Hampshire Public Health Labs. (NH)
New Hanover Regional Medical Center (NC)
New Lexington Clinic (KY)
New London Hospital (NH)
New Medical Centre Hospital (United Arab Emirates)
New York City Department of Health and Mental Hygiene (NY)
New York Eye and Ear Infirmary (NY)
New York Presbyterian Hospital (NY)
New York State Dept. of Health (NY)
New York University Medical Center (NY)
New Zealand Blood Service (New Zealand)
Newark Beth Israel Medical Center (NJ)
Newborn Metabloc Screening Program/ Alberta Health Services (Canada)
Newman Regional Health (KS)
Niagara Health System (Canada)
Ninewells Hospital and Medical School (United Kingdom [GB])
Noble's Hospital (United Kingdom [GB])
NorDx - Scarborough Campus (ME)
Norman Regional Hospital (OK)
North Carolina Baptist Hospital (NC)
North District Hospital (China)
North Kansas City Hospital (MO)
North Mississippi Medical Center (MS)
North Oaks Medical Center (LA)
North Philadelphia Health System-St. Joseph's Hospital (PA)
North Shore Hospital Laboratory (New Zealand)
North Shore Medical Center (MA)
North Shore-Long Island Jewish Health System Laboratories (NY)
North Vista Hospital (NV)
North York General Hospital (Canada)
Northcentral Technical College (WI)
Northcrest Medical Center (TN)
Northeast Georgia Health System (GA)
Northeastern Vermont Regional Hospital (VT)
Northern Virginia Community College (VA)
Northridge Hospital Medical Center (CA)
Northside Hospital (GA)
Northside Medical Center (OH)
Northumberland Hills Hospital (Canada)
Northwest Arkansas Pathology Associates (AR)
Northwestern Medical Center, Inc. (VT)
Northwestern Memorial Hospital (IL)
Norton Healthcare (KY)
Norwalk Hospital (CT)
Notre Dame Hospital (Canada)
Nova Scotia Association of Clinical Laboratory Managers (Canada)
Nova Scotia Community College (Canada)
Novus Path Labs (India)
NSW Health Pathology (Australia)
NW Physicians Lab (WA)
Oakton Community College (IL)
Ocean County Medical Laboratories (NJ)
Ochsner Clinic Foundation (LA)
Oconee Memorial Hospital (SC)
Odense University Hospital (Denmark)
Office of Medical Services Laboratory (DC)
Ohio Health Laboratory Services (OH)
Ohio State University Hospitals (OH)
Ohio Valley Medical Center (WV)
Oklahoma Heart Hospital, LLC (OK)
Oklahoma State University: Center for Health Sciences (OK)
Olive View-UCLA Medical Center (CA)
Olmsted Medical Center Laboratory (MN)
Ontario Agency for Health Protection and Promotion (Canada)
Ontario Medical Association Quality Management Program-Laboratory Service (Canada)
Onze Lieve Vrouwziekenhuis (Belgium)
Orange County Community College (NY)
Orange Park Medical Center (FL)
Ordre Professionnel Des Technologistes Médicaux Du Québec (Canada)
Orebro University Hospital (Sweden)
Oregon Public Health Laboratory (OR)
Orillia Soldiers Memorial Hospital (Canada)
Orlando Health (FL)
OSF - Saint Anthony Medical Center (IL)
Oslo University Hospital (Norway)
OSU Veterinary Diagnostic Laboratory (OR)
Ottawa Regional Hospital & Healthcare Center (IL)
OU Medical Center (OK)
Our Lady of the Lake Regional Medical Center/FMOL Health System (LA)
Our Lady's Hospital for Sick Children (Ireland)
Overlake Hospital Medical Center (WA)
Ozarks Medical Center (MO)
PA Veterinary Laboratory (PA)
Pacific Diagnostic Laboratories (CA)
Palmer Lutheran Health Center (IA)
Palmetto Baptist Medical Center (SC)
Palmetto Health Baptist Easley (SC)
Palo Alto Medical Foundation (CA)
Pamela Youde Nethersole Eastern Hospital (Hong Kong East Cluster) (Hong Kong)
Paris Community Hospital (IL)
Parkview Adventist Medical Center (ME)
Parkview Health Laboratories (IN)
Parkwest Medical Center (TN)
Parrish Medical Center (FL)
Pathgroup (TN)
Pathlab (IA)
Pathology Associates Medical Lab. (WA)
Pathology Resource Network (LA)
PathWest Laboratory Medicine WA (Australia)
PeaceHealth Laboratories (OR)
Peninsula Regional Medical Center (MD)
Penn State Hershey Medical Center (PA)
Pennsylvania Dept. of Health (PA)
Pennsylvania Hospital (PA)
Peoria Tazewell Pathology Group, P.C. (IL)
PEPFAR Tanzania (PA)
PerkinElmer Health Sciences, Inc. (SC)
Peterborough Regional Health Centre (Canada)
Peterson Regional Medical Center (TX)
PHIA Project, NER (CO)
Phoebe Putney Memorial Hospital (GA)
Phoenix Children's Hospital (AZ)
Phoenixville Hospital (PA)
Physicians Choice Laboratory Services (NC)
Physicians Laboratory & SouthEast Community College (NE)
Physicians Regional Medical Center (FL)
Piedmont Atlanta Hospital (GA)
Piedmont Henry Hospital (GA)
Pioneers Memorial Health Care District (CA)
Placer County Public Health Laboratory (CA)
Plains Memorial Hospital (TX)
Pocono Medical Center School of Medical Technology (PA)
Pointe Coupee Parish Hospital (LA)
Pomona Valley Hospital Medical Center (CA)
Portneuf Medical Center (ID)
Poudre Valley Hospital (CO)
Prairie Lakes Hospital (SD)
Presbyterian Hospital - Laboratory (NC)
Presbyterian/St. Luke's Medical Center (CO)
Preventive Medicine Foundation (Taiwan)
Prince George Regional Hospital (Canada)
Prince of Wales Hospital (Hong Kong)
Princess Margaret Hospital (Hong Kong)
Proasecal LTD (Colombia)
ProMedica Laboratory (OH)
Prometheus Laboratories Inc. (CA)
Providence Alaska Medical Center (AK)
Providence Everett Medical Center (WA)
Providence Hospital (AL)
Providence St. Joseph Medical Center (CA)
Providence St. Mary Medical Center (WA)
Provista Diagnostics (AZ)
Public Health Laboratory (Dublin) (Ireland)
Puget Sound Blood Center (WA)
Pullman Regional Hospital (WA)
Queen Elizabeth Hospital (Canada)
Queen Elizabeth Hospital (China)
Queen Mary Hospital (Hong Kong)
Queensland Health Pathology Services (Australia)

Quest - A Society for Adult Support and Rehabilitation (Canada)
Quincy Medical Center (MA)
Quinte Healthcare Corp. - Belleville General Site (Canada)
Quintiles Laboratories, Ltd. (GA)
Ramathibodi Hospital (Thailand)
Randers Regional Hospital (Denmark)
Range Regional Health Services (MN)
Ransom Memorial Hospital (KS)
Rapides Regional Medical Center (LA)
Rappahannock General Hospital (VA)
RCPA Quality Assurance Programs Pty Limited (Australia)
Reading Hospital (PA)
Regina Qu'Appelle Health Region (Canada)
Regional Laboratory of Public Health (Netherlands)
Regional Medical Laboratory, Inc. (OK)
Regions Hospital (MN)
Rehoboth McKinley Christian Health Care Services (NM)
Reid Hospital & Health Care Services (IN)
Renown Regional Medical Center (NV)
Research Institute of Tropical Medicine (Philippines)
Rhode Island Dept. of Health Labs (RI)
Rhode Island Hospital (RI)
Rice Memorial Hospital (MN)
Ridgeview Medical Center (MN)
Riverside Community Hospital (CA)
Riverside Health System (VA)
Riverside Medical Center (IL)
Riverside Medical Center (WI)
Riverton Memorial Hospital (WY)
Riverview Hospital (IN)
Riyadh Armed Forces Hospital, Sulaymainia (Saudi Arabia)
RMIT University (Australia)
Robert E. Bush Naval Hospital (CA)
Rochester General Hospital (NY)
Rockford Memorial Hospital (IL)
Roger Williams Medical Center (RI)
Roosevelt General Hospital (NM)
Roper St. Francis Healthcare (SC)
Ross University School of Veterinary Medicine (Saint Kitts and Nevis)
Roswell Park Cancer Institute (NY)
Rouge Valley Health System (Canada)
Round Rock Medical Center (TX)
Royal Children's Hospital (Australia)
Royal Hobart Hospital (Australia)
Royal Hospital (Oman)
Royal Melbourne Hospital (Australia)
Royal Victoria Hospital (Canada)
Rush University Medical Center (IL)
Russellville Hospital (AL)
SA Pathology (Australia)
SAAD Specialist Hospital (Saudi Arabia)
Sacred Heart Hospital (FL)
Sacred Heart Hospital (WI)
Sacred Hearth -St. Mary's Hospital Inc (WI)
Saddleback Memorial Medical Center (CA)
Sahlgrenska Universitetssjukhuset (Sweden)
Saint Francis Hospital & Medical Center (CT)
Saint Francis Medical Center (IL)
Saint Mary's Regional Medical Center (NV)
Salem Hospital (OR)
Salisbury University (MD)
Salzburger Landeskliniken (SALK) (Austria)
Samaritan Health Services (OR)
Samaritan Regional Health System (OH)
Samkwang Medical Laboratory (Korea, Republic of)
Sampson Regional Medical Center (NC)
Samsung Medical Center (Korea, Republic of)
San Angelo Community Medical Center (TX)
San Francisco General Hospital-University of California San Francisco (CA)
San Joaquin Community Hospital (CA)
San Jose State University (CA)
San Juan Regional Medical Group (NM)
Sanford Health (ND)
Sanford USD Medical Center (SD)
Santa Clara Valley Health & Hospital Systems (CA)
Santa Rosa Medical Center (FL)
Santiam Memorial Hospital (OR)
Sarasota Memorial Hospital (FL)
Saratoga Hospital (NY)
SARL Laboratoire Caron (France)
Saskatchewan Disease Control Laboratory (Canada)
Saskatoon Health Region (Canada)
Saudi Aramco Medical (TX)
SC Department of Health and Environmental Control (SC)
Schneck Medical Center (IN)
School of Animal and Veterinary Science, University of Adelaide (Australia)
Schuyler Hospital (NY)
Scientific Institute of Public Health (Belgium)
Scott & White Memorial Hospital (TX)
Scott Air Force Base (IL)
Scripps Health (CA)
Scuola Di Specializzaaione-University Milano Bicocca (Italy)
Seattle Cancer Care Alliance (WA)
Seattle Children's Hospital/Children's Hospital and Regional Medical Center (WA)
Sel Lam Terral (France)
Seminole Hospital District (TX)
Sentinel CH SpA (Italy)
Seoul National University Hospital (Korea, Republic of)
Seoul St. Mary's Hospital (Korea, Republic of)
Seton Healthcare Network (TX)
Seton Medical Center (CA)
Shands At the University of Florida (FL)
Shands Jacksonville (FL)
Shared Hospital Laboratory (Canada)
Sharon Regional Health System (PA)
Sharp Health Care Laboratory Services (CA)
Shiel Medical Laboratory Inc. (NY)
Shore Memorial Hospital (NJ)
Shriners Hospitals for Children (OH)
Shriners Hospitals for Children (SC)
Silliman Medical Center (Philippines)
Silverton Health (OR)
Sime Darby Medical Centre Subang Jaya Sdn. Bhd. (Malaysia)
SIMeL (Italy)
Singapore General Hospital (Singapore)
Singulex (CA)
Sky Lakes Medical Center (OR)
Slidell Memorial Hospital (LA)
Slotervaart Ziekenhuis (Netherlands)
SMDC Clinical Laboratory (MN)
Sociedad Espanola de Bioquimica Clinica y Patologia Molec. (Spain)
Sociedade Brasileira de Analises Clinicas (Brazil)
Sociedade Brasileira de Patologia Clinica (Brazil)
South Bay Hospital (FL)
South Bend Medical Foundation (IN)
South County Hospital (RI)
South Dakota State Health Laboratory (SD)
South Eastern Area Laboratory Services (Australia)
South Miami Hospital (FL)
South Peninsula Hospital (AK)
South Texas Laboratory (TX)
Southeast Alabama Medical Center (AL)
SouthEast Alaska Regional Health Consortium (SEARHC) (AK)
Southern Community Laboratories (New Zealand)
Southern Health Care Network (Australia)
Southern Hills Medical Center (TN)
Southern Maryland Hospital (MD)
Southern Pathology Services, Inc. (PR)
Southlake Regional Health Center (Canada)
Southwest General Health Center (OH)
Southwest Healthcare System (CA)
Southwestern Regional Medical Center (OK)
Sparks Health System (AR)
Sparrow Hospital (PA)
Spaulding Hospital Cambridge (MA)
Speare Memorial Hospital (NH)
Specialty Vet Path (WA)
Spectra East (NJ)
Spryfield Family Medical Center (Canada)
St Elizabeth Hospital (WI)
St Rose Dominican Hospital (NV)
St. Agnes Healthcare (MD)
St. Anthony Hospital (OK)
St. Antonius Ziekenhuis (Netherlands)
St. Barnabas Medical Center (NJ)
St. Charles Medical Center-Bend (OR)
St. Charles Parish Hospital (LA)
St. Clair Hospital (PA)
St. Croix Regional Medical Center (WI)
St. David's Medical Center (TX)
St. David's South Austin Hospital (TX)
St. Elizabeth Community Hospital (CA)
St. Elizabeth's Medical Center (NY)
St. Eustache Hospital (Canada)
St. Francis Health Center (CA)
St. Francis Hospital (MO)
St. Francis Hospital (SC)
St. Francis Hospital & Health Centers (NY)
St. John Hospital and Medical Center (MI)
St. John Medical Center (OH)
St. John's Hospital (IL)
St. John's Hospital & Health Center (CA)
St. John's Mercy Medical Center (MO)
St. John's Regional Health Center (MO)
St. Joseph Health Center (MO)
St. Joseph Hospital (CA)
St. Joseph Hospital (NH)
St. Joseph Medical Center (TX)
St. Joseph Regional Health Center (TX)
St. Joseph's Health Centre (Canada)
St. Joseph's Hospital & Medical Center (AZ)
St. Jude Children's Research Hospital (TN)
St. Jude Medical Center (CA)
St. Luke's Episcopal Hospital (TX)
St. Luke's Hospital (IA)
St. Luke's Hospital (MN)
St. Luke's Hospital (MO)
St. Luke's Hospital (PA)
St. Luke's Hospital at The Vintage (TX)
St. Luke's Medical Center (AZ)
St. Luke's Regional Medical Center (ID)
St. Luke's Treasure Valley Regional Medical Center (ID)
St. Mark's Hospital (UT)
St. Mary Medical Center (CA)
St. Mary Medical Center (PA)
St. Mary's Good Samaritan (IL)
St. Mary's Health Center (MO)
St. Mary's Hospital (MT)
St. Mary's Hospital (NJ)
St. Mary's Hospital (NY)
St. Mary's Hospital (WI)
St. Mary's Medical Center (IN)
St. Mary's Medical Center (WV)
St. Michael's Hospital (WI)
St. Nicholas Hospital (WI)
St. Olavs Hospital (Norway)
St. Peter's Bender Laboratory (NY)
St. Peter's Hospital (MT)
St. Rita's Medical Center (OH)
St. Tammany Parish Hospital (LA)
St. Thomas Hospital (TN)
St. Thomas-Elgin General Hospital (Canada)
St. Vincent Hospital (NM)
St. Vincent's Medical Center (FL)
Stanford Hospital and Clinics (CA)
State of Alabama (AL)
State of Ohio Corrections Medical Center Laboratory (OH)
State of Washington Public Health Labs (WA)
State of Wyoming Public Health Laboratory (WY)
Statens Serum Institut (Denmark)
Stillwater Medical Center (OK)
Stockton Pathology Medical Group (CA)
Stony Brook University Hospital (NY)
Stormont-Vail Regional Medical Ctr. (KS)
Sturgis Hospital (MI)
Sunnybrook Health Sciences Centre (Canada)
Sunrise Hospital and Medical Center (NV)
SUNY Downstate Medical Center (NY)
Supratech Micropath Lab & Research Institute (India)
Susan B. Allen Hospital (KS)
Susquehanna Health System (PA)
Sutter Health Sacramento Sierra Region Laboratories (CA)
Swedish American Health System (IL)
Swedish Medical Center (CO)
Sydney South West Pathology Service Liverpool Hospital (Australia)
Tahoe Forest Hospital (CA)
Taichung Veterans General Hospital (Taiwan)
Taiwan Society of Laboratory Medicine (Taiwan)
Tampa General Hospital (FL)
Tan Tock Seng Hospital (Singapore)
Taranaki Medlab (New Zealand)
Tartu University Clinics (Estonia)
Tataa Biocenter (Sweden)
Taylor Regional Hospital (KY)
Temple Community Hospital (CA)
Temple University Hospital - Parkinson Pavilion (PA)
Tenet Healthcare (PA)
Tethys Bioscience, Inc. (CA)
Tewksbury Hospital (MA)
Texas A & M University (TX)
Texas Children's Hospital (TX)
Texas Department of State Health Services (TX)
Texas Health Harris Methodist Hospital Cleburne (TX)
Texas Health Harris Methodist Hospital Fort Worth (TX)
Texas Health Presbyterian Hospital Dallas (TX)
Texas Scottish Rite Hospital for Children (TX)
The AGA Khan University Hospital (Pakistan)
The Broad Institute (MA)
The Brooklyn Hospital Center (NY)
The Charlotte Hungerford Hospital (CT)
The Cheshire Medical Center (NH)
The Children's Mercy Hospital (MO)
The City Hospital Dubai UAE (United Arab Emirates)
The Clinical Microbiology Institute (OR)
The Cooley Dickinson Hospital, Inc. (MA)
The First Hospital of China Medical University (China)
The Good Samaritan Hospital (PA)
The Hospital for Sick Children (Canada)
The Joint Commission (IL)
The Korean Society for Laboratory Medicine (Korea, Republic of)
The Michener Inst. for Applied Health Sciences (Canada)
The Naval Hospital of Jacksonville (FL)
The Nebraska Medical Center (NE)
The Norwegian Institute of Biomedical Science (Norway)

The Ohio State University-Vet Hospital (OH)
The Permanente Medical Group, Inc. (CA)
The University of Texas M.D. Anderson Cancer Center (TX)
The University of Texas Medical Branch (TX)
The University of the West Indies, Trinidad Campus (Trinidad and Tobago)
The University of Tokyo (Japan)
Thibodaux Regional Medical Center (LA)
Thomas Jefferson University Hospital, Inc. (PA)
Thunder Bay Regional Health Sciences Centre (Canada)
Torrance Memorial Medical Center (CA)
Touro Infirmary (LA)
Tri-Cities Laboratory (WA)
TriCore Reference Laboratories (NM)
Trident Medical Center (SC)
Trillium Health Partners Credit Valley Hospital (Canada)
Trinity Health Systems (OH)
Trinity Hospital of Augusta (GA)
Trinity Medical Center (AL)
Trinity Muscatine (IA)
Tripler Army Medical Center (HI)
Trumbull Memorial Hospital (OH)
Tucson Medical Center (AZ)
Tuen Mun Hospital, Hospital Authority (Hong Kong)
Tufts Medical Center Hospital (MA)
Tulane Medical Center Hospital & Clinic (LA)
Tulane University Health Sciences Center (LA)
Twin Lakes Regional Medical Center (KY)
Tyrone Hospital (PA)
U.S. Medical Ctr. for Federal Prisoners (MO)
U.S. Naval Hospital, Yokosuka, Japan (AP)
UC Davis Medical Center Department of Pathology & Laboratory Medicine (CA)
UC San Diego Health System Clinical Laboratories (CA)
UCI Medical Center (CA)
UCLA Medical Center (CA)
UCONN Health Center (CT)
UCSF Medical Center China Basin (CA)
UMass Memorial Medical Center (MA)
UMC of El Paso- Laboratory (TX)
UMC of Southern Nevada (NV)
Umea University Hospital (Sweden)
UNC Hospitals (NC)
Unidad De Patologia Clinica (Mexico)
Union Clinical Laboratory (Taiwan)
United Christian Hospital (Hong Kong)
United Clinical Laboratories (IA)
United Health Services Hospital/Wilson Hospital Lab (NY)
United Memorial Med Center (NY)
United States Air Force School of Aerospace Medicine/PHE (OH)
Universidade Federal Do Rio de Janeiro (Brazil)
Universitaet Zuerich (Switzerland)
Universitair Ziekenhuis Antwerpen (Belgium)
University College Hospital (Ireland)
University Health Network Laboratory Medicine Program (Canada)
University Hospital (GA)
University Hospital (TX)
University Hospital Center Sherbrooke (CHUS) (Canada)
University Hospitals of Cleveland (OH)
University Malaya Medical Centre (Malaysia)
University Medical Center (TN)
University Medical Center (TX)
University Medical Center at Lafayette (LA)
University Medical Center at Princeton (NJ)
University Medical Center of El Paso (TX)
University Medical Center Utrecht (Netherlands)
University of Alabama Hospital Lab (AL)
University of Alberta - Medical Genetics (Canada)
University of Arizona Medical Center (AZ)
University of Arkansas for Medical Sciences (AR)
University of Bonn (Germany)
University of British Columbia (Canada)
University of California Veterinary Medical Teaching Hospital (CA)
University of Chicago Hospitals Laboratories (IL)
University of Cincinnati Medical Center (OH)
University of Cologne Medical Center (Germany)
University of Colorado Health Sciences Center (CO)
University of Colorado Hospital (CO)
University of Connecticut (CT)
University of Delaware (DE)
University of Guadelajara Chemistry Department (Mexico)
University of Guelph (Canada)
University of Hong Kong (Hong Kong)
University of Idaho (ID)
University of Illinois Medical Center (IL)
University of Iowa Hospitals and Clinics (IA)
University of Iowa, Hygienic Lab (IA)
University of Kentucky Medical Center Hospital (KY)
University of Ljubljana Faculty of Medicine (Slovenia)
University of Louisville Hospital (KY)
University of Maryland Medical System (MD)
University of Miami (FL)
University of Miami - Clinical Genetics Labs (FL)
University of Minnesota Medical Center-Fairview (MN)
University of Missouri Hospital (MO)
University of MS Medical Center (MS)
University of New Mexico (NM)
University of North Carolina - Health Services (NC)
University of North Texas Health Science Center (TX)
University of Oregon (OR)
University of Pennsylvania (PA)
University of Pennsylvania Health System (PA)
University of Pittsburgh Medical Center (PA)
University of Portsmouth (United Kingdom [GB])
University of Queensland (Australia)
University of South Alabama Medical Center (AL)
University of Tasmania (Australia)
University of Tennessee, College of Veterinary Medicine (TN)
University of Texas Health Center (TX)
University of the Ryukyus (Japan)
University of TX M.D. Anderson Cancer Ctr. (TX)
University of Utah Hospital & Clinics (UT)
University of Virginia Medical Center (VA)
University of Washington Medical Center (WA)
University of Wisconsin Medical Foundation (WI)
UPMC Bedford Memorial (PA)
Urology of Indiana (IN)
Urology of Virginia, PLLC (VA)
USA MEDDAC-Japan
UT Southwestern Medical Center (TX)
Uvalde Memorial Hospital (TX)
UW Health (WI)
UZ-KUL Medical Center (Belgium)
VA (Alexandria) Medical Center (LA)
VA (Asheville) Medical Center (NC)
VA (Bay Pines) Medical Center (FL)
VA (Castle Point) Hudson Valley Health Care System (NY)
VA (Central Texas) Veterans Health Care System (TX)
VA (Chillicothe) Medical Center (OH)
VA (Columbus) (OH)
VA (Dayton) Medical Center (OH)
VA (Durham) Medical Center (NC)
VA (Grand Junction) Medical Center (CO)
VA (Huntington) Medical Center (WV)
VA (Indianapolis) Medical Center (IN)
VA (Miami) Medical Center (FL)
VA (Milwaukee) Medical Center (WI)
VA (Roseburg) Medical Center (OR)
VA (Tampa) Hospital (FL)
VA (Tuscaloosa) Medical Center (AL)
Vail Valley Medical Center (CO)
Valley Health/Winchester Medical Center (VA)
Valley Medical Center (WA)
Vancouver Island Health Authority (SI) (Canada)
Vanderbilt University Medical Center (TN)
Vejle Hospital (Denmark)
Vermont Department of Health (VT)
Vernon Memorial Hospital (WI)
Veterans Memorial Hospital (IA)
Via Christi Regional Medical Center (KS)
Virginia Commonwealth University (VA)
Virginia Mason Medical Center (WA)
Virginia Physicians, Inc. (VA)
Virginia Regional Medical Center (MN)
Virtua - West Jersey Hospital (NJ)
Wabash General Hospital (IL)
WakeMed (NC)
Walter Reed Army Institute of Research (MD)
Warren Hospital (NJ)
Washington Hospital Center (DC)
Waterbury Hospital (CT)
Waters Technologies Ireland Ltd (Ireland)
Watertown Memorial Hospital (WI)
Watson Clinic (FL)
Waukesha Memorial Hospital (WI)
Wayne Memorial Hospital (GA)
Weber State University (UT)
Weed Army Community Hospital Laboratory (CA)
Weeneebayko General Hospital (Canada)
Weirton Medical Center (WV)
Wellington Regional Medical Center (FL)
Wellstar Douglas Hospital Laboratory (GA)
Wellstar Health Systems (GA)
WellStar Paulding Hospital (GA)
Wenatchee Valley Medical Center (WA)
Wesley Medical Center (KS)
West Georgia Health Systems (GA)
West Penn Allegheny Health System-Allegheny General Hospital (PA)
West Shore Medical Center (MI)
West Valley Medical Center Laboratory (ID)
West Virginia Bureau for Public Health (WV)
West Virginia Univ. Hospitals (WV)
Westchester Medical Center (NY)
Western Baptist Hospital (KY)
Western Healthcare Corporation (Canada)
Western Nebraska Community College (NE)
Western State Hospital (VA)
Whangarei Hospital (New Zealand)
Wheaton Franciscan Laboratories At St. Francis (WI)
Wheeling Hospital (WV)
White Memorial Medical Center (CA)
Whitehorse General Hospital (Canada)
Whitman Hospital & Medical Center (WA)
Wickenburg Community Hospital (AZ)
William Beaumont Army Medical Center (TX)
William Beaumont Hospital (MI)
William Osler Health Centre (Canada)
Williamson Medical Center (TN)
Wilson Medical Center (NC)
Winchester Hospital (MA)
Winn Army Community Hospital (GA)
Winter Haven Hospital, Inc. (FL)
Wisconsin State Laboratory of Hygiene (WI)
Wishard Health Sciences (IN)
Womack Army Medical Center (NC)
Women & Infants Hospital (RI)
Womens and Childrens Hospital (LA)
Women's Health Care Group of PA (PA)
Woods Memorial Hospital (TN)
Woodside Health Center (Canada)
Wyckoff Heights Medical Center (NY)
Wyoming County Community Hospital (NY)
Yale New Haven Hospital (CT)
Yale University School of Medicine (CT)
York Hospital (PA)
Yukon-Kuskokwim Delta Regional Hospital (AK)
Yuma Regional Medical Center (AZ)
Zhongshan Hospital Fudan University (China)
Zuni PHS Indian Hospital (NM)

Individuals

Erika B Ammirati (CA)
Stephen Apfelroth (NY)
Deborah Bishop (WV)
Abbejane Blair (MA)
Vanessa Buchan (New Zealand)
A. Bjoern Carle (ME)
Tony Chan (China)
Omer Eltoum (Qatar)
Sahar Gamil EL-Wakil (Saudi Arabia)
Mary Lou Gantzer (DE)
Carlos Gonzalez (TX)
Natalie J. Kennel (CA)
Dr. Muain Haseeb (Saudi Arabia)
Judy Horton (MD)
B. Y. Hsieh (Taiwan)
Clark B Inderlied (CA)
Ellis Jacobs (NJ)
Nilesh Shah (CA)
Harvey Ronald Kennedy, MD (NJ)
Natalie J. Kennel (CA)
William F. Koch (MD)
Jan Krouwer (MA)
Jennifer Kwon (NY)
Debra Larsen (TX)
Sarah B Leppanen (CA)
Stefano A. Lollai (Italy)
Roberta Madej (CA)
Laura Miller (CA)
Samir Osman (Qatar)
A. K Peer (South Africa)
Armando Perez-Cardona (FL)
Jing Zhang (CA)
C. Anne Pontius (TN)
Aida Porras (Colombia)
Philip A Poston, PhD (FL)
Tawni Reller (MN)
Lisa Reninger (IL)
Nilesh Shah (CA)
Dinah Shore Myers (NC)
Abdullah Mohd. Siddiqi (Saudi Arabia)
Judi Smith (MD)
David Soloy (TX)
Steffini Stalos (TX)
Luke Thiboutot (MA)
Suresh H Vazirani (India)
Matthew A Wikler (NJ)
Jing Zhang (CA)

CLINICAL AND LABORATORY STANDARDS INSTITUTE®

Explore the Latest Offerings from CLSI!

As we continue to set the global standard for quality in laboratory testing, we're adding initiatives to bring even more value to our members and customers.

THE Key TO Quality™
from the Clinical and Laboratory Standards Institute (CLSI)

Power Forward with this Official Interactive Guide

Fundamentals for implementing a quality management system in the clinical laboratory.

Find Membership Opportunities

See the options that make it even easier for your organization to take full advantage of CLSI benefits and our unique membership value.

CLSI U
Laboratory Standards Education Center

Visit the CLSI U Education Center

Where we provide the convenient and cost-effective education resources that laboratories need to put CLSI standards into practice, including webinars, workshops, and more.

Shop Our Online Products

Including eCLIPSE Ultimate Access™, CLSI's cloud-based, online portal that makes it easy to access our standards and guidelines—*anytime, anywhere.*

For more information, visit www.clsi.org today.

CLINICAL AND LABORATORY STANDARDS INSTITUTE®

950 West Valley Road, Suite 2500, Wayne, PA 19087 USA

P: 610.688.0100 **Toll Free (US):** 877.447.1888 **F:** 610.688.0700

E: customerservice@clsi.org **www.clsi.org**

PRINT ISBN 1-56238-887-8

ELECTRONIC ISBN 1-56238-888-6