CLINICAL AND LABORATORY STANDARDS INSTITUTE®

3rd Edition

EP09c

Measurement Procedure Comparison and Bias Estimation Using Patient Samples

This guideline covers 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 medical 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 medical 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 advances 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 managed according to the consensus process by a committee of experts.

Appeal Process

When it is believed that an objection has not been adequately considered and responded to, the process for appeal, documented in the CLSI *Standards Development Policies and Processes,* is followed.

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 additional 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: +1.610.688.0100
F: +1.610.688.0700
www.clsi.org
standard@clsi.org

EP09c, 3rd ed.
June 2018
Replaces EP09-A2-IR and EP09-A3

Measurement Procedure Comparison and Bias Estimation Using Patient Samples

Jeffrey R. Budd, PhD
A. Paul Durham, MA
Thomas E. Gwise, PhD
Douglas M. Hawkins, PhD
Mark Holland, 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 guideline EP09—*Measurement Procedure Comparison and Bias Estimation Using Patient Samples* 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 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 chapter 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*. 3rd ed. CLSI guideline EP09c (ISBN 978-1-68440-006-5 [Print]; ISBN 978-1-68440-007-2 [Electronic]). Clinical and Laboratory Standards Institute, 950 West Valley Road, Suite 2500, Wayne, Pennsylvania 19087 USA, 2018.

CLINICAL AND LABORATORY STANDARDS INSTITUTE®

Suggested Citation*

CLSI. *Measurement Procedure Comparison and Bias Estimation Using Patient Samples*. 3rd ed. CLSI guideline EP09c. Wayne, PA: Clinical and Laboratory Standards Institute; 2018.

Previous Editions:
January 1986, April 1993, December 1995, September 2002, July 2010, August 2013

ISBN 978-1-68440-006-5 (Print)
ISBN 978-1-68440-007-2 (Electronic)
ISSN 1558-6502 (Print)
ISSN 2162-2914 (Electronic) Volume 38, Number 12

* EP09 was corrected in June 2018, and the code was revised to EP09c.

Committee Membership

Consensus Council

Dennis J. Ernst, MT(ASCP), NCPT(NCCT)
Chairholder
Center for Phlebotomy Education
USA

Mary Lou Gantzer, PhD, FACB
Vice-Chairholder
USA

J. Rex Astles, PhD, FACB, DABCC
Centers for Disease Control and Prevention
USA

Lucia M. Berte, MA, MT(ASCP)SBB, DLM, CQA(ASQ)CMQ/OE
Laboratories Made Better!
USA

Karen W. Dyer, MT(ASCP), DLM
Centers for Medicare & Medicaid Services
USA

Thomas R. Fritsche, MD, PhD, FCAP, FIDSA
Marshfield Clinic
USA

Loralie J. Langman, PhD, DABCC, FACB, F-ABFT
Mayo Clinic
USA

Ross J. Molinaro, PhD, MLS(ASCP)CM, DABCC, FACB
Siemens Healthcare Diagnostics, Inc.
USA

James R. Petisce, PhD
BD Diagnostic Systems
USA

Andrew Quintenz
Bio-Rad Laboratories, Inc.
USA

Robert Rej, PhD
New York State Department of Health – Wadsworth Center
USA

Zivana Tezak, PhD
FDA Center for Devices and Radiological Health
USA

Expert Panel on Evaluation Protocols

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

Paula Ladwig, MS, MT(ASCP)
Vice-Chairholder
Mayo Clinic
USA

Valeria L. Alcon, PhD
Health Canada
Canada

Julianne Cook Botelho, PhD
Centers for Disease Control and Prevention
National Center for Environmental Health
USA

Jeffrey R. Budd, PhD
Beckman Coulter
USA

Mark D. Kellogg, PhD, MT(ASCP), DABCC, NACB
Boston Children's Hospital
USA

Marina V. Kondratovich, PhD
FDA Center for Devices and Radiological Health
USA

Robert J. McEnroe, PhD
USA

Jeffrey E. Vaks, PhD
Roche Molecular Diagnostics
USA

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

Jeffrey R. Budd, PhD
Chairholder
Beckman Coulter
USA

A. Paul Durham, MA
USA

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

Anders Kallner, MD, PhD
Karolinska Hospital
Stockholm, Sweden

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

Staff

Clinical and Laboratory Standards Institute
USA

Luann Ochs, MS
Project Manager

Megan L. Tertel, MA, ELS
Editorial Manager

Catherine E.M. Jenkins
Editor

Kristy L. Leirer, MS
Editor

Laura Martin
Editor

Acknowledgment

CLSI, the Consensus Council, the Expert Panel on Evaluation Protocols, and the Document Development Committee on Method Comparison and Bias Estimation Using Patient Samples gratefully acknowledge the following individuals for their thorough review of this guideline and the resulting corrections and clarifications reflected in the corrected edition:

Jeffrey R. Budd, PhD
Chairholder
Beckman Coulter
USA

Douglas M. Hawkins, PhD
University of Minnesota Medical Center - Fairview
USA

A. Paul Durham, MA
USA

Mark Holland, PhD
Beckman Coulter
USA

Contents

Abstract i

Committee Membership iii

Foreword vii

Chapter 1: Introduction 1

1.1 Scope 1
1.2 Standard Precautions 1
1.3 Terminology 2

Chapter 2: Introduction to Measurement Procedure Comparison and Bias Estimation 7

2.1 Process Flow Chart 7
2.2 Introduction 8

Chapter 3: Preparing for the Measurement Procedure Comparison Study 13

3.1 Planning for Measurement Procedure Comparison Studies 13
3.2 Considerations for the Medical Laboratory 18

Chapter 4: Performing the Measurement Procedure Comparison Study 19

4.1 Measurement Procedure Familiarization Period 19
4.2 Quality Control 19
4.3 Running the Study 20
4.4 Inspection of Data During Collection 20
4.5 Documentation of Rejected Data 20

Chapter 5: Visual Data Analysis 21

5.1 Visual Data Review 21
5.2 Scatter Plots 21
5.3 Difference Plots 22
5.4 Inspect Plots for Underlying Characteristics 23

Chapter 6: Quantitative Analysis 31

6.1 Estimating Bias From Difference Plots 31
6.2 Fitting a Line to Scatter Plots (Regression Analysis) 36
6.3 Bias and Regression Parameters With Confidence Intervals 41

Chapter 7: Comparisons Within a Measurement Procedure 43

7.1 Comparing Two Conditions Within an Already Validated Measurement Procedure 43
7.2 Sample Type Comparisons 43
7.3 Other Comparisons 43

Chapter 8: Interpreting Results and Comparing to Performance Criteria 45

8.1 Interpreting the Study Results 45
8.2 Manufacturer's Statement of Bias Performance Claims 46
8.3 Medical Laboratory's Statement of Bias Performance 47

Chapter 9: Conclusion 48

Contents (Continued)

Chapter 10: Supplemental Information 48

References 49

Appendix A. Deming Regression 51

Appendix B. Weighted Deming Regression 56

Appendix C. A Practical Example Illustrating Bias Estimation and Measurement Procedure Comparison Techniques 59

Appendix D. Example Datasets 67

Appendix E. Detecting Aberrant Results (Outliers) 71

Appendix F. Confidence Interval of a Median Estimate of Bias Between Measurement Procedures 74

Appendix G. Ordinary Linear Regression 80

Appendix H. Weighted Least Squares Regression (Weighted Ordinary Linear Regression) 83

Appendix I. Passing-Bablok Regression 90

Appendix J. Profile Weighted Deming Regression 93

Appendix K. Iterative Approaches for Estimating Confidence Intervals for Bias and Regression Parameters 96

The Quality Management System Approach 102

Related CLSI Reference Materials 103

Foreword

Measurement procedure comparison is one of the most common techniques used by both manufacturers and medical 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, measurand concentration) 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 medical 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 validation of performance claims for bias. In the case of a laboratorian, it involves introducing a measurement procedure into the laboratory, including verification of its manufacturer's claims (specifications). The scope of the experimental and data-handling procedures for these two purposes differs. 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 most appropriate measurement procedure currently available. The laboratorian usually compares the candidate and an available procedure. Then, there might 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 laboratory 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 guideline is intended to promote effective and correct data analysis and reporting using standard experimental and statistical methods.

Manufacturers of medical laboratory measurement procedures or devices should use this guideline 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.

Overview of Changes

This guideline replaces the previous edition of the approved guideline, EP09-A2-IR, published in 2010. Several changes were made in this edition, including:

- Broader coverage of measurement procedure 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

This guideline was corrected in 2018 and replaces the original third edition of the approved guideline, EP09-A3, published in 2013. Corrections were made as follows:

- Reorganizing the content to emphasize the process of performing a measurement procedure comparison
- Clearly specifying that manufacturers should use regression analysis to characterize bias
- Adding information on using precision profile information in performing Deming regressions
- Adding more information on determining confidence intervals for bias estimates at specified concentrations using regression fits
- Making corrections to the description of the Passing-Bablok regression technique
- Adding a detailed description of the bootstrap iterative technique for bias estimation
- Correcting minor miscellaneous errors in equations

> **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.

NOTE: The content of this guideline is supported by the CLSI consensus process and does not necessarily reflect the views of any single individual or organization.

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

Chapter 1: Introduction

This chapter includes:

- Guideline's scope and applicable exclusions
- Standard precautions information
- "Note on Terminology" that highlights particular use and/or variation in use of terms and/or definitions
- Terms and definitions used in the guideline
- Abbreviations and acronyms used in the guideline
- Symbols used in the guideline

1.1 Scope

This guideline provides recommendations 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 both difference plot and regression procedures to determine the relationship 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 authorities and medical laboratory personnel.

This guideline is for use with measurement procedures that provide quantitative numerical results. This guideline is not intended for use with ordinal IVD examinations, commonly referred to as qualitative procedures (see CLSI document EP12[1]). This guideline is not intended to provide information on 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]).

1.2 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 bloodborne pathogens. Published guidelines are available that discuss the daily operations of diagnostic medicine in humans and animals while encouraging a culture of safety in the laboratory.[6] 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.[7]