Evaluation of Precision Performance of Quantitative Measurement Methods; Approved Guideline—Second Edition

Daniel W. Tholen, M.S.
Anders Kallner, M.D., Ph.D.
John W. Kennedy (deceased)
Jan S. Krouwer, Ph.D.
Kristen Meier, Ph.D.

Abstract

NCCLS document EP5-A2, Evaluation of Precision Performance of Quantitative Measurement Methods; Approved Guideline—Second Edition provides guidance and procedures for evaluating the precision of in vitro diagnostic devices and includes recommendations for manufacturers in evaluating their devices and methods when establishing performance claims. Included are guidelines for the duration, procedures, materials, data summaries, and interpretation techniques that are adaptable for the widest possible range of analytes and device complexity. The procedures are designed for manufacturers or developers of clinical laboratory measurement methods, and for users of those methods who wish to determine their own performance capabilities or to verify claims from a manufacturer. A balance is created in the document between complexity of design and formulae, and simplicity of operation.

Contents

Abstract ....................................................................................................................................................i

Committee Membership ........................................................................................................................ iii

Foreword .............................................................................................................................................. vii

1 Scope ..................................................................................................................................................1

2 Introduction ....................................................................................................................................... .1

3 Standard Precautions ........................................................................................................................1

4 Definitions ..........................................................................................................................................1

5 Symbols Used in Text ........................................................................................................................3

6 Overview of the General Precision Evaluation Experiment ..............................................................4

   6.1 General Guidelines ..........................................................................................................................4

   6.2 Device Familiarization Period ........................................................................................................4

   6.3 Protocol Familiarization Period .......................................................................................................4

   6.4 Precision Evaluation Experiment ..................................................................................................4

   6.5 Completing the Precision Experiment ..........................................................................................5

   6.6 Comparison to Other Precision Evaluations ................................................................................5

7 Statistical Power of Precision Estimates ..............................................................................................5

   7.1 Precision and Confidence ................................................................................................................5

   7.2 Statistical Comparison with the Manufacturer ..............................................................................6

8 Device Familiarization Period .............................................................................................................6

   8.1 Purpose .........................................................................................................................................6

   8.2 Duration ........................................................................................................................................6

9 Protocol Familiarization Period (for Users and Manufacturers) ............................................................6

   9.1 Purpose .........................................................................................................................................6

   9.2 Duration ........................................................................................................................................7

   9.3 Use of Data ....................................................................................................................................7

   9.4 Quality Control Procedures ..........................................................................................................7

   9.5 Additional Evaluations ..................................................................................................................7

   9.6 Preliminary Precision Evaluation ..................................................................................................7

10 Precision Evaluation Experiment ......................................................................................................7

   10.1 Components of Precision ...............................................................................................................8

   10.2 Reagents and Calibration Materials ............................................................................................8

   10.3 Test Materials .............................................................................................................................8

   10.4 Number of Runs and Days ..........................................................................................................9

   10.5 Recording the Data ......................................................................................................................10

   10.6 Quality Control Procedures ........................................................................................................10

   10.7 Detection of Outliers ..................................................................................................................11

   10.8 Statistical Calculations for Precision ..........................................................................................11

   10.9 Comparison with Manufacturers’ Claims or Other Performance Criteria ..................................13
Contents (Continued)

11 Use of These Guidelines by Manufacturers to Establish Precision Performance ...............15
    11.1 Factors to be Considered ...................................................................................15
    11.2 Incorporating Multiple Factors .........................................................................16
    11.3 Format for Statement of Claims .......................................................................16

References .......................................................................................................................19

Additional References ........................................................................................................20

Appendix A. Sample Data Recording Sheets ..................................................................21

Appendix B. Example of Completed Sample Data Recording Sheets ..............................25

Appendix C. Additional Statistical Considerations ........................................................29

Summary of Consensus/Delegate Comments and Committee Responses ......................35

The Quality System Approach .........................................................................................38

Related NCCLS Publications ..........................................................................................39
Foreword

Current clinical chemistry literature contains numerous examples of product evaluations. Many of these use the basic concepts that are included in this guideline. While more complex and customized experimental designs have been used for both published studies and regulatory purposes in special cases, there still appears to be a strong need in the clinical community for the basic approaches to quantitative precision assessment to be described, as well as their rationales.

In order to address this need, the committee has drawn on the experience of users, representatives of industry, statisticians, chemists, laboratory personnel, regulatory authorities, and medical personnel for developing this guideline. The extremely wide variety of in vitro diagnostic devices currently available made it apparent that a single experimental design would not be appropriate for all devices. Therefore, this guideline has been constructed to provide primarily conceptual guidance on the duration, procedures, materials, data summaries, and interpretation techniques that would be adaptable for the widest possible range of analytes and device complexity. Illustrations of each step of the evaluation, with an example of a typical experimental design, have also been provided.

In development of this protocol, many recommendations for duration, inclusion of quality control, and methods of determining the components of precision were carefully considered. The resultant protocol creates a balance between complexity of design and formulae, and simplicity of operation. For ease of use, an appendix (Appendix C) has been included that provides guidelines for modifying the design and calculations when appropriate.

Since its publication as a tentative guideline in 1992 and then as an approved guideline in 1999, document EP5 has been widely used by device manufacturers to establish precision claims for their methods, and by laboratories to determine the performance of methods in their use. In preparing EP5-A2, the Area Committee on Evaluation Protocols sought to retain the experimental procedures while harmonizing the terminology with current international recommendations. Along with terminology changes, text has been added to distinguish different users of the document, and sections were added on Scope and Definitions, per current NCCLS document format. Comments received on the present edition are included in the comment/response summary starting on page 35. The area committee plans to broaden the scope of the next revision of this document to address more complex procedural issues and expand the statistical analysis involved in evaluating precision.

The various components of precision in EP5-A2, especially “within-laboratory precision,” could be important components of measurement uncertainty (MU), per current ISO usage: “a parameter, associated with the result of a measurement that characterizes the dispersion of the values that could reasonably be attributed to the measurand.” However, EP5-A2 procedures cannot be used alone to estimate MU. There are other components for many analytes, and different ways to combine them. Fully ISO-compliant calculations of measurement uncertainty involve concepts and procedures that are beyond the scope of the current document, but may be addressed in the next edition.

A Note on Terminology

NCCLS, as a global leader in standardization, is firmly committed to achieving global harmonization in terminology wherever possible. Harmonization is a process of recognizing, understanding, and explaining differences in terms while taking steps to achieve worldwide uniformity. NCCLS 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 NCCLS, ISO, and CEN documents; and that legally required use of terms, regional usage, and different consensus timelines are all obstacles to harmonization. In light of this, NCCLS recognizes that harmonization of terms facilitates the global application of standards and deserves immediate attention. Implementation of this policy must be an evolutionary and educational process that begins with new projects and revisions of existing documents.
In order to align the usage of terminology in this document with that of ISO, the term *sample* has replaced the term *specimen* and the term *measuring range* has replaced the term *reportable range*. The users of EP5-A2 should understand that the fundamental meanings of the terms are similar, and to facilitate understanding, where appropriate, the terms are defined along with their ISO counterpart in the guideline’s Definitions section. (See Section 4.)

The term *precision* is always used as a measure of “closeness of agreement between independent test/measurement results obtained under stipulated conditions.”1 The terms in this document are consistent with uses defined in the ISO 3534 and ISO 5725 series of standards. In these models, *repeatability* and *reproducibility* are considered to be the extreme measures of precision, with repeatability being the smallest measure (same operator, method, equipment, time, and laboratory) and reproducibility being the largest (different operator, equipment, and laboratory). All other measures of precision are “intermediate measures” and must be explicitly described. Reproducibility is not estimated in EP5-A2, since the protocol does not require multiple laboratories. All other measures of precision from EP5-A have been retained, although the term *total precision* was eliminated, because it was not clearly defined. In this document, *total precision* has been replaced by *within-laboratory* or *within-device*, depending on whether the laboratory or manufacturer is deriving the estimate.

**Key Words**

Evaluation protocol, experimental design, medical devices, outlier, precision, quality control
Evaluation of Precision Performance of Quantitative Measurement Methods; Approved Guideline—Second Edition

1 Scope

This document provides guidelines for experiments to evaluate the precision performance characteristics of quantitative measurement methods and devices. It includes protocols for developers of testing methods or devices, and protocols for users of these methods who wish to determine their own performance capabilities. These procedures may not be appropriate for some quantitative methods for which adequate test materials do not exist.

2 Introduction

This document is for manufacturers of in vitro diagnostic (IVD) devices and developers of clinical laboratory measurement methods who wish to establish the precision capabilities of their methods. It is also for the users of those methods who wish to verify the validity of performance claims, or who simply want to measure their own precision. Users of automated measurement procedures who wish only to apply a minimal protocol to verify the validity of a manufacturer’s claims for precision should follow the guidance of the most current edition of NCCLS document EP15—User Demonstration of Performance for Precision and Accuracy. The guidelines are fully general for these situations, because they include considerations of goals for the reliability of the precision estimates.

This document also applies to laboratories that make significant modifications to current methods. When using a modification of an in vitro diagnostic (IVD) device or method, a user needs to verify that essential performance characteristics of the device have not changed. Comparison to original claimed precision performance may not be valid. Examples of typical modifications are the use of reagents, sample sources, calibrating or control materials, or operating procedures that are different from those stated in the manufacturer’s labeling (instructions for use).

3 Standard Precautions

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

4 Definitions

Analyte – Component represented in the name of a measurable quantity; NOTES: a) This includes any element, ion, compound, substance, factor, infectious agent, cell, organelle, activity (enzymatic, hormonal, or immunological), or property, the presence or absence, concentrations, activity, intensity, or other characteristics of which are to be determined; b) In the type of quantity “mass of protein in 24-hour