



# EP12-A2

## User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline—Second Edition

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.

# 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 Administrative Procedures.

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](http://www.clsi.org)  
[standard@clsi.org](mailto:standard@clsi.org)

ISBN 1-56238-654-9  
ISSN 0273-3099

EP12-A2  
Vol. 28 No. 3  
Replaces EP12-A  
Vol. 22 No. 14

---

## User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline—Second Edition

Volume 28 Number 3

Patricia E. Garrett, PhD  
Fred D. Lasky, PhD  
Kristen L. Meier, PhD

### Abstract

Clinical and Laboratory Standards Institute document EP12-A2—*User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline—Second Edition* provides the user with a consistent approach for protocol design and data analysis when evaluating qualitative diagnostic tests. Guidance is provided for both precision and method-comparison studies.

Clinical and Laboratory Standards Institute (CLSI). *User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline—Second Edition*. CLSI document EP12-A2 (ISBN 1-56238-654-9). Clinical and Laboratory Standards Institute, 950 West Valley Road, Suite 2500, Wayne, Pennsylvania 19087 USA, 2008.

The Clinical and Laboratory Standards Institute consensus process, which is the mechanism for moving a document through two or more levels of review by the health care community, is an ongoing process. Users should expect revised editions of any given document. Because rapid changes in technology may affect the procedures, methods, and protocols in a standard or guideline, users should replace outdated editions with the current editions of CLSI documents. Current editions are listed in the CLSI catalog and posted on our website at [www.clsi.org](http://www.clsi.org). If your organization is not a member and would like to become one, and to request a copy of the catalog, contact us at: Telephone: 610.688.0100; Fax: 610.688.0700; E-Mail: [customerservice@clsi.org](mailto:customerservice@clsi.org); Website: [www.clsi.org](http://www.clsi.org).



Copyright ©2008 Clinical and Laboratory Standards Institute. Except as stated below, any reproduction of content from a CLSI copyrighted standard, guideline, companion product, or other material requires express written consent from CLSI. All rights reserved. Interested parties may send permission requests to [permissions@clsi.org](mailto:permissions@clsi.org).

CLSI hereby grants permission to each individual member or purchaser to make a single reproduction of this publication for use in its laboratory procedure manual at a single site. To request permission to use this publication in any other manner, e-mail [permissions@clsi.org](mailto:permissions@clsi.org).

### **Suggested Citation**

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.

### **Proposed Guideline**

July 2000

### **Approved Guideline**

August 2002

### **Approved Guideline—Second Edition**

January 2008

ISBN 1-56238-654-9  
ISSN 0273-3099

## Contents

Abstract.....	i
Committee Membership.....	iii
Foreword.....	vii
Laboratory Error Sources and CLSI Evaluation Protocols Documents.....	ix
1 Scope.....	1
2 Introduction.....	1
3 Standard Precautions.....	2
4 Clinical Utility.....	2
4.1 Screening Tests.....	3
4.2 Diagnostic Tests.....	3
4.3 Confirmatory Tests.....	3
5 Terminology.....	3
5.1 A Note on Terminology.....	3
5.2 Terms for Comparative Benchmark.....	4
5.3 Definitions.....	5
6 Device Familiarization and Training.....	8
6.1 Purpose.....	8
6.2 Duration.....	8
7 Evaluation Materials.....	8
7.1 Controls.....	8
7.2 Specimen Collection and Handling.....	8
8 Bias and Imprecision Studies.....	8
8.1 Analyte Concentrations Near the Cutoff.....	9
8.2 Negative and Positive Controls.....	11
8.3 A Qualitative Method-Precision Experiment for Analyte Concentrations Near the $C_{50}$ .....	12
9 Comparison of Methods.....	14
9.1 Test Specimens.....	15
9.2 Number of Specimens.....	16
9.3 Duration.....	16
9.4 Inspection of Data During Collection.....	16
9.5 Discrepant Results.....	16
9.6 Reference Specimen Panels.....	17
10 Data Analysis.....	17
10.1 Comparator Is Diagnostic Accuracy Criteria.....	17
10.2 Comparator Other Than Diagnostic Accuracy Criteria.....	23
10.3 Examples.....	25
References.....	32
Appendix. Statistical Reasoning for Precision Experiment Conclusions.....	34
Summary of Consensus Comments and Working Group Responses.....	40

**Contents (Continued)**

Summary of Delegate Comments and Working Group Responses .....41  
The Quality Management System Approach .....44  
Evaluation Protocols Documents, Descriptions, and Key Words.....45

SAMPLE

# User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline—Second Edition

## 1 Scope

This guideline provides protocols for the evaluation of qualitative test performance characteristics. In this document, a qualitative test is restricted to those tests that have only two possible outcomes (eg, positive/negative, present/absent, reactive/nonreactive). EP12 is written primarily for individuals and laboratories that use and evaluate such tests. These protocols are intended to help users determine test performance in their own testing environment. This guideline for qualitative test performance evaluation should help the device developer and the user to meet documentation and regulatory goals. While this document is not intended for manufacturers to establish test performance characteristics, the data analysis principles described here can be used by manufacturers.

Test methods with values that are reported as, for instance, negative, +1, +2, or +3, or as endpoint dilutions (commonly in multiples of 8, reflecting the microtiter plate format, or in multiples of 10) are often called semiquantitative. These methods are not further discussed in EP12, although if one of the values results is considered the cutoff for a positive test, the evaluation protocol recommended here could be applied to that cutoff. For instance, if a test for antibodies to the Lyme disease pathogen was reported as positive if the endpoint titer was  $\geq 1:160$ , the precision and method comparison experiments discussed below could be applied to that cutoff.

## 2 Introduction

Qualitative tests return one of two possible responses. Method evaluation procedures for such tests are diverse, with each laboratory specialty often emphasizing different issues in the experimental design, data analysis, and interpretation of such studies. EP12 offers a defined approach to method evaluation for many qualitative tests.

Clinical laboratories develop and implement qualitative tests for a number of reasons. Laboratories should document that the test performs as intended in their facilities, by operators who are expected to use the device. Often, such demonstration is required by laboratory regulatory or accreditation bodies.

Qualitative candidate methods are diverse, employing technologies from lateral flow with visual reading, to automated nucleic acid sequencing and base calling, to microarrays. While universal evaluation guidelines may not be feasible, common features exist. For example, precision studies and comparison of methods studies with patient specimens can be used to demonstrate each type of test's performance capabilities.

In addition to technical diversity, qualitative tests may differ just by having a different cutoff or medical decision point. For example, a qualitative candidate method for screening blood donations for an infectious disease, such as human immunodeficiency virus (HIV) or hepatitis B virus (HBV), would have a cutoff chosen to ensure a high sensitivity and high negative predictive value (NPV), so there is a correspondingly high probability that an infected unit would be excluded from the blood supply. The same method could be used to make a qualitative test for diagnosis, with a cutoff that is chosen to minimize the number of false-positive and false-negative results.

When a qualitative test is used as an aid in the diagnosis of an infectious disease, such as HIV, Streptococcus, or Trichomonas, the "cutoff" is the "limit of blank." (See the Definitions section and CLSI/NCCLS document EP17.<sup>5</sup>) The objective is to recognize the presence of the characteristic, such as immunologic response in the host or the presence of an antigen of the pathogen, which is indicative of

infection. In these cases, recording a measurable response from the device for a “negative” result is often not possible because, by definition, “negative” means the characteristic is below the “limit of blank.” Additionally, for many microbiological qualitative tests (such as the type often called “rapid tests”), there is no recognized reference material. For many such antigen or antibody tests, response as a function of “concentration” is not feasible. In those cases, cutoff information, although useful, might not be consistent between sample types (serum vs urine or a swab vs a saline wash), detected organisms (when several organisms have a common antigen), immune response, or device batches. Users must then rely on the labeling provided by the manufacturer for specific information on how the cutoff or limit of blank was established; or gain experience with the device in their patient population to correlate test results with clinical findings.

When the examination is performed by operator observation only, without the objective response generally provided by instrumentation, the variability caused by operator interpretation can be large relative to the uniformity of the device within and between manufacturing lots.<sup>1</sup> In such cases, variability is highly dependent upon operators’ abilities to detect the presence of color, which is affected by light conditions, background color (for contrast), individual judgment, and other parameters.

A protocol is provided that can be used to judge performance as a function of an independently controlled parameter, such as dilution (or titer, the inverse of dilution), with small numbers of operators. The dependent variable is the ratio of accurate readings (seeing color when the characteristic is present above the cutoff and not seeing color when it is absent) compared to total readings from several operators using replicates of the same sample. The comparison of performance between candidate and comparison methods can be heavily dependent upon the source and nature of the characteristic detected. In addition, the use of contrived materials (dilutions or commercial controls) as test samples might be advantageous for consistency of preparation, but might not accurately represent patient samples.

Some nucleic acid testing devices further illustrate the complexity of qualitative test evaluations with exquisitely sensitive techniques that can detect single strands of DNA or RNA, which is not necessarily the same as detecting disease.<sup>6</sup> When using or evaluating a diagnostic test, the meaning of a positive or negative result in the context of that test method must be clear.

The objective of this document is to provide guidance for performing uniform, well-defined studies that can be used to adequately evaluate and describe performance characteristics of qualitative tests.

### **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 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 US Centers for Disease Control and Prevention.<sup>7</sup> 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 CLSI document M29.<sup>8</sup>

### **4 Clinical Utility**

Qualitative tests may be used for a wide range of clinical purposes. Their uses can be described as screening, diagnostic, confirmatory, or monitoring. The test’s sensitivity and specificity, predictive values, and the prevalence of the disease or condition in the population tested determine the clinical utility of the qualitative test.



### 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 approach is based on the model presented in the most current edition of CLSI/NCCLS document HS1—*A Quality Management System Model for Health Care*. The quality management system approach applies a core set of “quality system essentials” (QSEs), basic to any organization, to all operations in any 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:

- Documents & Records Organization Personnel
- Equipment Purchasing & Inventory Process Control
- Information Management Occurrence Management Assessments—External & Internal
- Process Improvement Customer Service Facilities & Safety

EP12-A2 addresses the QSEs indicated by an “X.” For a description of the other documents listed in the grid, please refer to the Related CLSI Reference Materials section on the following page.

Documents & Records	Organization	Personnel	Equipment	Purchasing & Inventory	Process Control	Information Management	Occurrence Management	Assessments—External & Internal	Process Improvement	Customer Service	Facilities & Safety
	EP18	EP18			X C24 EP5 EP6 EP7 EP9 EP10 EP14 EP15 EP17 EP18 EP19 EP21 GP10 M29		EP18	EP10 EP18	EP7 EP18		M29

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

## Evaluation Protocols Documents, Descriptions, and Key Words\*

- EP5-A2** **Evaluation of Precision Performance of Quantitative Measurement Methods; Approved Guideline—Second Edition (2004).** This document provides guidance for designing an experiment to evaluate the precision performance of quantitative measurement methods; recommendations on comparing the resulting precision estimates with manufacturers' precision performance claims and determining when such comparisons are valid; as well as manufacturers' guidelines for establishing claims.
- Evaluation protocol, experimental design, medical devices, outlier, precision, quality control
- EP6-A** **Evaluation of the Linearity of Quantitative Measurement Procedures: A Statistical Approach; Approved Guideline (2003).** This document provides guidance for characterizing the linearity of a method during a method evaluation; for checking linearity as part of routine quality assurance; and for determining and stating a manufacturer's claim for linear range.
- Allowable difference, allowable error, linearity, matrix effects, measurement error, total error, uncertainty
- EP7-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.
- Evaluation, hazard analysis, interference, interferent, matrix effects, performance claims, risk management, specificity, validation, verification
- EP9-A2** **Method Comparison and Bias Estimation Using Patient Samples; Approved Guideline—Second Edition (2002).** This document addresses procedures for determining the bias between two clinical methods, and the design of a method comparison experiment using split patient samples and data analysis.
- Bias, evaluation protocol, experimental design, linear regression, method comparison, quality control, residuals
- EP10-A3** **Preliminary Evaluation of Quantitative Clinical Laboratory Measurement Procedures; Approved Guideline—Third Edition (2006).** This guideline provides experimental design and data analysis for preliminary evaluation of the performance of a measurement procedure or device.
- Carry-over, comparison of methods, drift, evaluation protocol, experimental design, linearity, multiple regression, outlier, precision
- 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.
- Analytical interference, bias, matrix, matrix effect, physicochemical interference
- EP15-A2** **User Verification of Performance for Precision and Trueness; Approved Guideline—Second Edition (2005).** 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.
- Bias, precision, repeatability, trueness, verification of performance
- EP17-A** **Protocols for Determination of Limits of Detection and Limits of Quantitation; Approved Guideline (2004).** This document provides guidance for determining the lower limit of detection of clinical laboratory methods, for verifying claimed limits, and for the proper use and interpretation of the limits.
- Limit of blank, limit of detection, limit of quantitation, nonparametric statistics

---

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

## Evaluation Protocols Documents, Descriptions, and Key Words (Continued)

**EP18-A**     **Quality Management for Unit-Use Testing; Approved Guideline (2002).** This guideline recommends a quality management system for unit-use devices that will aid in the identification, understanding, and management of sources of error (potential failure modes) and help to ensure correct results. It is targeted for those involved in the supervision of laboratory-testing quality management, and it addresses issues related to specimen collection through reporting of test results.

Quality assurance, quality control, quality management, quality system, unit-use system

**EP19-R**     **A Framework for NCCLS Evaluation Protocols; A Report (2002).** This document describes the different types of performance studies that are conducted to evaluate clinical assays.

Demonstration, evaluation protocol, validation, verification

**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 which 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.

Error, error of measurement, measurement error, total analytical error, total analytical error interval

**GP10-A**     **Assessment of the Clinical Accuracy of Laboratory Tests Using Receiver Operating Characteristics (ROC) Plots; Approved Guideline (1995).** This document provides a protocol for evaluating the accuracy of a test to discriminate between two subclasses of subjects where there is some clinically relevant reason to separate them. In addition to the use of ROC plots, the importance of defining the question, selecting the sample group, and determining the "true" clinical state are emphasized.

Clinical accuracy, false-negative fraction, false-positive fraction, medical decision analysis, performance evaluation, receiver operating characteristic (ROC) plot, sensitivity, specificity, true-negative fraction, true-positive fraction

### Other Related Publications

**C24-A3**     **Statistical Quality Control for Quantitative Measurement Procedures: Principles and Definitions; Approved Guideline—Third Edition (2006).** This guideline provides definitions of analytical intervals; plans for quality control procedures; and guidance for quality control applications.

**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.



# 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.



## 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.



## 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](http://www.clsi.org) today.

SAMPLE



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](mailto:customerservice@clsi.org) [www.clsi.org](http://www.clsi.org)

ISBN 1-56238-654-9