



CLINICAL AND
LABORATORY
STANDARDS
INSTITUTE®

1st Edition

EP39

A Hierarchical Approach to Selecting Surrogate Samples for the Evaluation of *In Vitro* Medical Laboratory Tests

This guideline establishes a definition of a surrogate sample, provides recommendations for determining when to use surrogate samples, and describes a process for selecting the most appropriate surrogate sample.

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

P: +1.610.688.0100

F: +1.610.688.0700

www.clsi.org

standard@clsi.org

A Hierarchical Approach to Selecting Surrogate Samples for the Evaluation of *In Vitro* Medical Laboratory Tests

April Veoukas, BS, JD

Shannon Bennett, MS, MBA, CMQOE(ASQ)

Melissa Barhoover, PhD, RAC

Natalya Benina, MS

Marvin Berman, PhD

Marc D. Goldford, BS

Shiaolan Y. Ho, PhD

Mark D. Kellogg, PhD, MT(ASCP), DABCC, FAACC

Mary Knighten, MT(AMT)

Marina V. Kondratovich, PhD

Gregory Jay Pomper, MD

Paula Ladwig, MS, MT(ASCP)

Fred D. Lasky, PhD

Sangeetha Vijaysri Nair, DVM, PhD

Patrick O'Donnell, BA

Michael Pikulski, PhD, DABCC, FACC

Yvonne Shea, MS

Bin Zhang, MD, MS

Abstract

Clinical and Laboratory Standards Institute guideline EP39—*A Hierarchical Approach to Selecting Surrogate Samples for the Evaluation of In Vitro Medical Laboratory Tests* establishes a standard definition of a surrogate sample. It presents a hierarchical approach for determining when to use surrogate samples and selecting an appropriate one. It also describes elements of a surrogate sample plan and includes technical preparation guidance for the characteristic to be measured or detected and for artificial matrix compositions. This guideline provides examples for specific performance study types.

Clinical and Laboratory Standards Institute (CLSI). *A Hierarchical Approach to Selecting Surrogate Samples for the Evaluation of In Vitro Medical Laboratory Tests*. 1st ed. CLSI guideline EP39 (ISBN 978-1-68440-130-7 [Print]; ISBN 978-1-68440-131-4 [Electronic]). Clinical and Laboratory Standards Institute, USA, 2021.

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.

If you or your organization is not a member and would like to become one, or to request a copy of the catalog, contact us at:

P: +1.610.688.0100 **F:** +1.610.688.0700 **E:** customerservice@clsi.org **W:** www.clsi.org



CLINICAL AND
LABORATORY
STANDARDS
INSTITUTE®

Copyright ©2021 Clinical and Laboratory Standards Institute. Except as stated below, any reproduction of content from a CLSI copyrighted standard, guideline, derivative product, or other material requires express written consent from CLSI. All rights reserved. Interested parties may send permission requests to 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 procedures manual at a single site. To request permission to use this publication in any other manner, e-mail permissions@clsi.org.

Suggested Citation

CLSI. *A Hierarchical Approach to Selecting Surrogate Samples for the Evaluation of In Vitro Medical Laboratory Tests*. 1st ed. CLSI guideline EP39. Clinical and Laboratory Standards Institute; 2021.

EP39-Ed1

ISBN 978-1-68440-130-7 (Print)

ISBN 978-1-68440-131-4 (Electronic)

ISSN 1558-6502 (Print)

ISSN 2162-2914 (Electronic)

Volume 41, Number 21

Committee Membership

Consensus Council

James R. Petisce, PhD

Chairholder

BD Diagnostic Systems

USA

Avis Danishefsky, PhD

FDA Center for Devices and

Radiological Health

USA

M. Laura Parnas, PhD, DABCC

Roche Diagnostics

USA

Tania Motschman, MS, MT(ASCP)SBB

Vice-Chairholder

USA

Collette Fitzgerald, PhD

Centers for Disease Control and

Prevention

USA

Victoria Petrides, MS

Abbott

USA

Deirdre Astin, MS, MT(ASCP)

USA

Michelle McLean, MS, MT(ASCP), BS

Greiner Bio-One, Inc.

USA

Matthew A. Wikler, MD, FIDSA, MBA

IDTD Consulting

USA

Anne T. Daley, MS, MT(ASCP)DLM,

CMQ/OE(ASQ)CSBB

USA

James H. Nichols, PhD, DABCC, FAACC

Vanderbilt University School of

Medicine

USA

Document Development Committee on Surrogate Sample Framework

April Veoukas, BS, JD

Chairholder

Abbott Laboratories

USA

Natalya Benina, MS

Fujirebio Diagnostics

USA

Mary Knighten, MT(AMT)

Diagnostica Stago

USA

Shannon Bennett, MS, MBA,

CMQOE(ASQ)

Vice-Chairholder

Mayo Clinic

USA

Cheryl Dobbe, MLT, MT(ASCP)

Centers for Medicare & Medicaid

Services

USA

Marina V. Kondratovich, PhD

FDA Center for Devices and

Radiological Health

USA

Mark D. Kellogg, PhD, MT(ASCP),

DABCC, FAACC

Boston Children's Hospital

USA

Gregory J. Pomper, MD

College of American Pathologists

USA

Expert Panel on Evaluation Protocols

Paula Ladwig, MS, MT(ASCP)

Chairholder

Mayo Clinic

USA

James H. Nichols, PhD, DABCC,

FAACB

Vice-Chairholder

Vanderbilt University School of

Medicine

USA

Valeria L. Alcon, PhD

Health Canada

Canada

J. Rex Astles, PhD, DABCC, FAACC

Centers for Disease Control and
Prevention

USA

Jeffrey R. Budd, PhD

USA

A. Paul Durham, MA

APD Consulting

USA

Brett Holmquist, PhD, ASCP, DABCC,

FAACC

LabCorp - Endocrine Sciences

USA

Jesper V. Johansen, PhD

Radiometer Medical A/S

Denmark

Edward Ki Yun Leung, PhD, DABCC,

FAACC

Children's Hospital Los Angeles

USA

Stephen Lovell, BS, PhD

FDA Center for Devices and

Radiological Health

USA

Nancy S. Miller, MD

Boston University School of Medicine

USA

Jeffrey E. Vaks, PhD

Roche Molecular Diagnostics

USA

Staff

Clinical and Laboratory Standards

Institute

USA

Laura Martin

Editorial Manager

Kristy L. Leirer, MS

Editor

Tabitha Kern, MS, MLS(ASCP)^{CM}

Project Manager

Catherine E.M. Jenkins, ELS

Editor

Acknowledgment

CLSI, the Consensus Council, and the Document Development Committee on Surrogate Sample Framework gratefully acknowledge the following volunteers for their important contributions to the development of this guideline:

Melissa Barhoover, PhD, RAC

Diagnostica Stago

USA

Fred D. Lasky, PhD

Lasky Consulting

Michael Pikulski, PhD, DABCC, FACC

Sonic Reference Laboratory

USA

Marvin Berman, PhD

Abbott Laboratories

USA

Sangeetha Vijaysri Nair, DVM, PhD

Hologic, Inc.

Yvonne Shea, MS

FDA Center for Devices and

Radiological Health

USA

Marc D. Goldford, BS

Sekisui Diagnostics

USA

Patrick O'Donnell, BA

Roche Molecular Diagnostics

Bin Zhang, MD, MS

Siemens-Healthineers

USA

Shiaolan Y. Ho, PhD

Abbott Laboratories

USA

Contents

Abstract	i
Committee Membership	iii
Foreword	vii
Chapter 1: Introduction	1
1.1 Scope	2
1.2 Background	2
1.3 Standard Precautions	3
1.4 Terminology	3
Chapter 2: Path of Workflow	9
Chapter 3: Samples	13
3.1 Patient Sample	14
3.2 Surrogate Sample	14
Chapter 4: Surrogate Sample Hierarchical Approach	17
4.1 Decision to Use Surrogate Samples	18
4.2 Surrogate Sample Hierarchy	18
4.3 Surrogate Sample Composition	19
Chapter 5: Surrogate Sample Plan	23
5.1 Identifying Needed Samples	25
5.2 Developing a Surrogate Sample Plan	26
Chapter 6: Technical Preparation Techniques	27
6.1 Overview	28
6.2 Determining the Analyte	28
6.3 Selecting the Matrix	30
6.4 Surrogate Sample Stability	30
6.5 Process Steps	31
6.6 Storage	31

Contents (Continued)

Chapter 7: Special Considerations.....	33
7.1 Qualifying the Surrogate Sample	34
7.2 Demonstrating Comparability	35
7.3 Using Surrogate Samples for More Than One Study	36
7.4 Preparing Blank Surrogate Samples	37
7.5 Creating a Single Signature Score From Multiple Genes	39
7.6 Rare Subtypes	40
7.7 Surrogate Samples for Molecular Assays.....	41
Chapter 8: Application to Performance Studies.....	43
8.1 Linearity.....	44
8.2 Analytical Specificity	46
8.3 Precision.....	48
8.4 Detection Capability.....	49
8.5 Matrix Comparison.....	52
8.6 Method Comparison	54
8.7 Hook Effect.....	58
8.8 Sample Stability	59
8.9 Carryover.....	61
8.10 Reagent Stability	62
Chapter 9: Conclusion	65
Chapter 10: Supplemental Information	67
References.....	68
Additional Resources.....	72
Appendix A. Surrogate Sample Descriptions and Examples	73
Appendix B. Example Chemical Compositions for Artificial Matrixes	76
Appendix C. <i>Chlamydia trachomatis</i> : Precision/Reproducibility	79
Appendix D. Creatinine: Precision/Reproducibility.....	101
Appendix E. Total Prostate Specific Antigen: Precision/Reproducibility	106
Appendix F. Albumin: Method Comparison	112
Appendix G. <i>Chlamydia trachomatis</i> : Instrument Carryover	117
The Quality Management System Approach.....	128
Related CLSI Reference Materials.....	129

Foreword

Terms such as “contrived,” “altered,” “processed,” “diluted,” “supplemented,” and “simulated” have been used interchangeably to describe substitutions for patient samples. This guideline establishes a uniform term, “surrogate sample,” and definition to describe material(s) that is used as a substitute for body fluid or tissue from a single human individual.

When appropriately characterized patient samples are unavailable, surrogate samples serve an important role in the development, validation, and verification of laboratory tests. Surrogate samples may be needed for many reasons, including limited sample volume or inadequate numbers of patient samples with concentrations at medical decision levels or at the extremes of the analytical measuring interval. A lack of available patient samples may be due to low disease prevalence, invasive sampling methods, or other reasons.

This guideline establishes an approach for selecting, preparing, and using surrogate samples. It describes the principles for creating a surrogate sample plan and presents a hierarchy, by performance study type, for selecting an appropriate surrogate sample. The hierarchical approach is demonstrated through product- and performance-specific examples.

KEY WORDS

Artificial analyte	Pooled	Supplemented
Artificial matrix	Sample plan	Surrogate sample
Hierarchy	Simulated	

This page is intentionally left blank.

Chapter 1

Introduction

This chapter includes:

- Guideline's scope and applicable exclusions
- Background information pertinent to the guideline's content
- Standard precautions information
- Terminology information, including:
 - Terms and definitions used in the guideline
 - Abbreviations and acronyms used in the guideline

A Hierarchical Approach to Selecting Surrogate Samples for the Evaluation of *In Vitro* Medical Laboratory Tests

1 Introduction

1.1 Scope

This guideline establishes a definition of “surrogate sample” and an approach for selecting, preparing, and using these samples. It discusses surrogate sample:

- Composition
- Technical preparation
- Selection criteria
- Documentation and planning
- Use in specific performance study types

The intended users of this guideline are *in vitro* diagnostic (IVD) device developers, laboratorians, and regulators. This guideline does not describe performance study design, which is covered in other standards and guidelines (see CLSI document EP19¹).

1.2 Background

Development, validation, and verification of laboratory tests depends on the availability of patient samples for testing. When appropriate patient samples are unavailable to validate test performance, using surrogate samples enables more efficient use of biological materials, improves testing efficiency, and facilitates the development of tests for new biomarkers. Patient samples for test development and other uses may be unavailable for several reasons.

Reasons that patient samples cannot be used include:

- Logistical constraints
- Insufficient sample volumes
- Inadequate numbers of samples, such as those with concentrations at medical decision levels (MDLs) or at the extremes of the analytical measuring interval (AMI)
- Technical constraints
- Unsatisfactory samples (ie, that lack the necessary characteristics for a performance study)
- Instability of samples
- Unavailability of blank or negative samples

Reasons that adequate patient specimens cannot be collected include:

- Low disease prevalence
- Limited target patient population
- Compromised patient health
- Invasive sampling methods
- Ethical considerations

The use of surrogate samples is not new. Surrogate samples have long played an important role in test development, validation, and verification. However, the lack of standard terminology and defined guidance for designing, selecting, preparing, and using surrogate samples reduces the predictability, timeliness, and efficiency of test development and the availability of tests for use.

Terms such as “contrived,” “altered,” “processed,” “diluted,” “supplemented,” and “simulated” are frequently used interchangeably with the term “surrogate.” Sometimes, they are considered subcategories of samples within the broader category of surrogate samples. The Document Development Committee on Surrogate Sample Framework reviewed relevant vocabularies and was unable to find a common definition of “surrogate sample.” Varying terminology for surrogate samples impedes the development of sound scientific strategies for their use. This guideline establishes uniform terminology and recommendations for the justifiable use of surrogate samples in test development, performance evaluations, and test methods.

1.3 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 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.² 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 M29.³

1.4 Terminology

CLSI, as a global leader in standardization, is firmly committed to achieving global harmonization whenever 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 different countries and regions and that legally required use of terms, regional usage, and different consensus timelines are all important considerations in the harmonization process. CLSI recognizes its important role in these efforts, and its consensus process focuses on harmonization of terms to facilitate the global application of standards and guidelines. Table 1 is provided to clarify the intended interpretations of the following terms.

Table 1. Common Terms or Phrases With Intended Interpretations

Term or Phrase	Intended Interpretation
“Needs to” or “must”	Explains an action directly related to fulfilling a regulatory and/or accreditation requirement or is indicative of a necessary step to ensure patient safety or proper fulfillment of a procedure
“Require”	Represents a statement that directly reflects a regulatory, accreditation, performance, product, or organizational requirement or a requirement or specification identified in an approved documentary standard
“Should”	Describes a recommendation provided in laboratory literature, a statement of good laboratory practice, or a suggestion for how to meet a requirement

1.4.1 Definitions

For purposes of this guideline, the terms and definitions listed below apply. Consult CLSI’s Harmonized Terminology Database at <https://htd.clsi.org> for related terms and definitions.

altered biological matrix – milieu in which the analyte exists that has been modified or changed through additional processes not routinely conducted as part of the normal specimen collection process (eg, treatment with heat or sodium hydroxide, clarification of the specimen milieu by centrifugation⁴).

analyte – component represented in the name of a measurable quantity⁵; **NOTE:** In the type of quantity “mass of protein in 24-hour urine,” “protein” is the analyte. In “amount of substance of glucose in plasma,” “glucose” is the analyte. In both cases, the long phrase represents the measurand.⁵

analytical measuring interval (AMI) – 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 conditions⁶; **NOTE 1:** The AMI is the concentration interval within which a given accuracy is achieved and a linear relationship between measurand values and the measurement procedure’s results is maintained with no preexamination dilution, concentration, or other pretreatment not part of the standard or routine measurement; **NOTE 2:** This interval is separate from the extended measuring interval, within which preexamination dilutions are needed to obtain results.

analytical specificity – ability of a measurement procedure to determine solely the quantity it purports to measure.⁷

artificial analyte – component created or manufactured for use in spiking surrogate samples, with the intention of emulating the characteristic of the human specimen that will be measured or detected⁴; **EXAMPLES:** Synthesized chemical compounds, plasmid or cell-line DNA, transcript RNA, recombinant protein, cell culture used to reproduce an artificially created organism.⁴

artificial matrix – milieu that is created or manufactured with the intention of emulating the specimen biological matrix⁴; **EXAMPLES:** Collection media, elution media, isotonic solutions, human or bovine serum albumin dissolved in buffer.⁴

biological analyte – analyte composed of or derived from a specimen for use in spiking surrogate samples⁴; **EXAMPLES:** Analyte-positive specimens; purified or extracted protein; DNA extracted from formalin-fixed, paraffin-embedded tissue specimens; analyte obtained from purified organisms; biological analyte reproduced from cell culture.⁴

biological matrix – components of a material system, except the analyte⁵; **NOTE:** The biological matrix is the milieu in which the analyte exists in a given sample (eg, serum, urine).

biological reference interval – specified interval of the distribution of values taken from a biological reference population.^{8,9}

commutability (of a material) – closeness of agreement between the mathematical relationship of the measurement results obtained by two measurement procedures for a stated quantity in a given material and the mathematical relationship obtained for the quantity in routine samples.¹⁰

interference (analytical) – artifactual increase or decrease in apparent concentration or intensity of an analyte (or measurand) due to the presence of a substance that reacts nonspecifically with either the detecting reagent or the signal itself.

in-use stability – duration of time over which the performance of a product within its expiration date remains within specified limits after the container system supplied by the manufacturer is opened and the product is put into use under standard operating conditions (eg, storage on the instrument).

limit of blank (LoB) – highest measurement result that is likely to be observed, with a stated probability (α), for a blank sample; **NOTE:** Typically, $\geq 95\%$ of all observed results on blank samples have concentration measurements less than the LoB when tested under routine medical laboratory conditions and in a defined type of sample.

linearity – ability (within a given interval) to provide results that are directly proportional to the concentration (or amount) of the analyte in the test sample.

lower limit of detection (LLoD) – measured quantity value, obtained by a given measurement procedure, for which the probability of falsely claiming the absence of a component in a material is β , given a probability α of falsely claiming its presence⁶; **NOTE:** In quantitative and qualitative molecular measurement procedures, the lowest concentration of analyte that can be consistently detected (typically, in $\geq 95\%$ of samples tested under routine medical laboratory conditions).

lower limit of quantitation (LLoQ) – lowest concentration of analyte in a sample that can be quantitatively determined with stated acceptable precision and trueness, under stated experimental conditions.

matrix (of a material system) – components of a material system, except the analyte.^{7,11}

matrix comparison – equivalence of matrix types.⁴

matrix effect – influence of a property of the sample, independent of the presence of the analyte, on the measurement and thereby on the measured quantity value.¹¹

measurand – quantity intended to be measured.⁶

method comparison – evaluation of bias or systematic difference between two quantitative tests or evaluation of agreements between results of two qualitative tests.⁴

patient sample – a specimen without modification or a specimen modified before measurement; **NOTE:** Examples of modifications are use of chemical additives, extraction, centrifugation, or other physical separation techniques and preexamination steps.

precision (of measurement) – closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions.⁶

qualitative tests – candidate methods that provide only two categorical responses (ie, positive/negative or yes/no).

repeatability (measurement) – measurement precision under a set of repeatability conditions of measurement.⁶

reproducibility (measurement) – measurement precision under reproducibility conditions of measurement⁶;

NOTE: The closeness of agreement between the results of the measurements of the same measurand performed under changed conditions of measurement.

sample – one or more parts taken from a system and intended to provide information about the system or to serve as a basis for a decision about the system⁷; **EXAMPLE:** A volume of serum taken from a larger volume of serum⁸; **NOTE:** For the purposes of EP39, a sample may be physically or chemically changed from the original patient specimen (see **specimen**), as in having been spiked with a potentially interfering substance.

shelf-life – period of time until the expiry date during which an *in vitro* diagnostic reagent in its original packaging maintains its stability under the storage conditions specified by the manufacturer.^{4,9}

specimen – discrete portion of a body fluid, breath, hair, or tissue taken for examination, study, or analysis of one or more quantities or properties assumed to apply for the whole.⁸

surrogate sample – material or combination of materials used as a substitute for body fluid or tissue taken for examination from a single human subject to study the characteristic of interest. Surrogate samples include but are not limited to⁴:

- Pooled patient samples of biological origin
- Materials supplemented (eg, spiked) with an analyte of interest
- Material created to have properties similar to or representative of the body fluid or tissue of interest
- Material composed of a combination of an analyte that simulates the analyte of interest and a matrix created to have properties similar to or representative of the body fluid or tissue or of the patient or subject
- More-complex combinations of fabricated analyte and matrix

unaltered biological matrix – milieu in which the analyte exists that has not been modified or changed from the normal specimen collection process⁴; **EXAMPLES:** Serum; plasma; urine; cerebrospinal fluid; swabs stored in collection media; formalin-fixed, paraffin-embedded tissue.⁴

1.4.2 Abbreviations and Acronyms

AMI analytical measuring interval

βhCG β-human chorionic gonadotropin

C5 the value in a relevant scale where a qualitative examination declares a sample to be positive 5% of the time

C95 the value in a relevant scale where a qualitative examination declares a sample to be positive 95% of the time

CMV cytomegalovirus

CSF cerebrospinal fluid

DNA deoxyribonucleic acid

EDTA ethylenediaminetetraacetic acid

ELISA	enzyme-linked immunosorbent assay
FFPE	formalin-fixed, paraffin-embedded
IVD	<i>in vitro</i> diagnostic
LLMI	lower limit of the measuring interval
LLoD	lower limit of detection
LLoQ	lower limit of quantitation
LoB	limit of blank
MDL	medical decision level
NPA	negative percent agreement
NSCLC	non–small cell lung cancer
pH	negative logarithm of hydrogen ion concentration
PPA	positive percent agreement
QC	quality control
RNA	ribonucleic acid
ULMI	upper limit of the measuring interval

This page is intentionally left blank.

Chapter 2

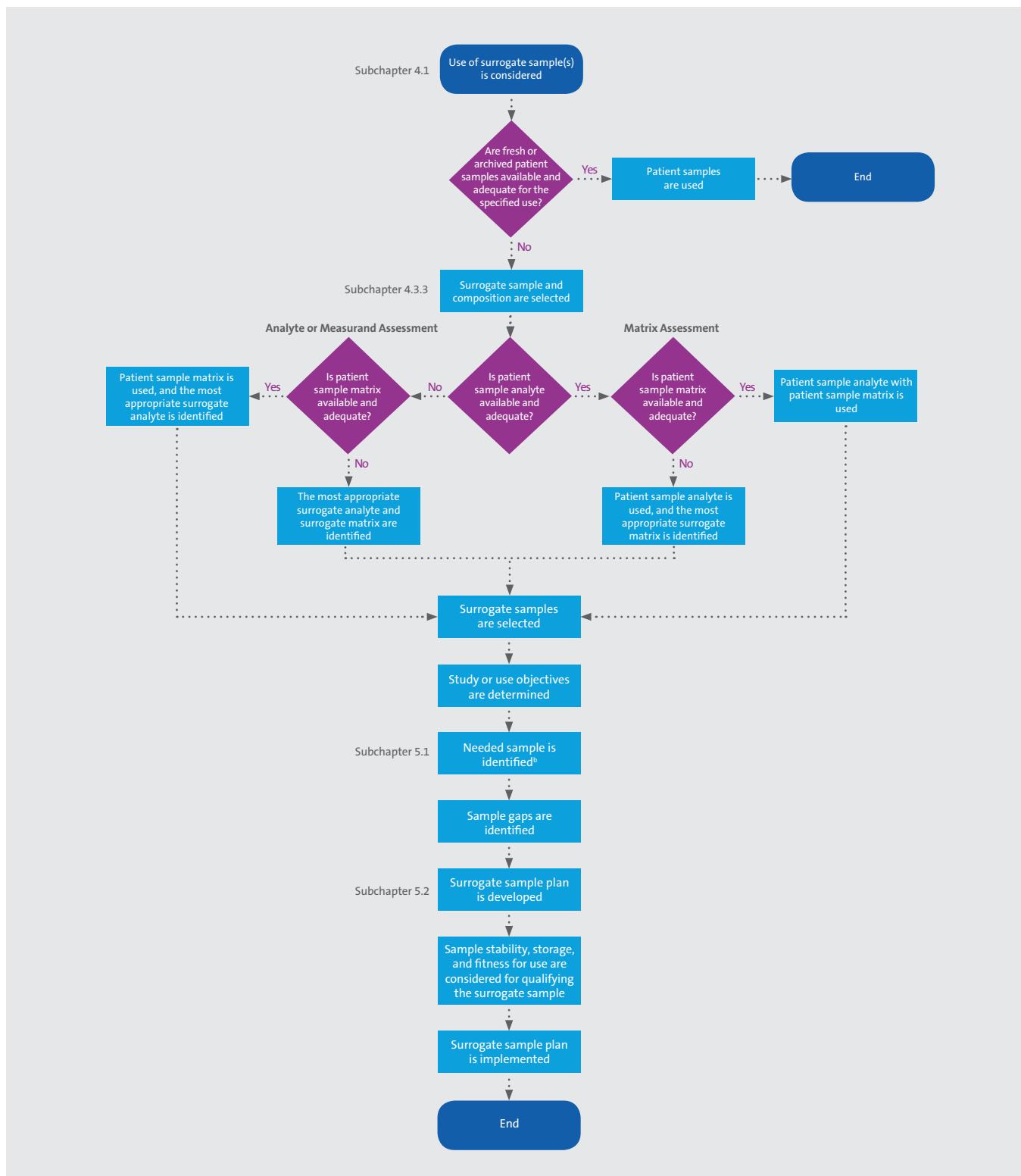
Path of Workflow

This chapter includes:

- A process flow chart for using surrogate samples

2 Path of Workflow

The decision to use a surrogate sample involves choosing the sample that is most appropriate for the type of study to be conducted. The process flow chart shown in Figure 1 begins with the consideration of whether to use a surrogate sample and shows the actions involved in each step of selecting a suitable surrogate sample and creating a surrogate sample plan.



^a Four basic symbols are used in this process flow chart: oval (signifies the beginning or end of a process), arrow (connects process activities), box (designates process activities), diamond (includes a question with alternative "Yes" and "No" responses).

^b The analyte, matrix, volume, and concentration should mimic those of the patient sample.

Figure 1. Process for Using Surrogate Samples^a

This page is intentionally left blank.

Chapter ③

Samples

This chapter includes:

- Patient sample description
- Surrogate sample definition
- Description of surrogate sample categories

3 Samples

3.1 Patient Sample

An analyte sampled for testing originates from the human body, where two components are in balance: the analyte and the surrounding natural matrix. The matrix of the patient sample is all sample components except the analyte. For example, when blood sodium concentrations are measured, sodium is the analyte, while everything else in the blood sample is the matrix. Physical separation methods (eg, centrifugation), use of standard chemical additives in collection devices (eg, EDTA- or heparin-coated tubes), and extraction techniques (eg, those specified in IVD device package inserts) that are necessary to test for the measurand do not alter the designation as a patient sample.

Good medical laboratory practice relies on test methods that were established using patient samples derived from sources that are as close to the intended specimen as possible. For example, in practice, blood and urine specimens are readily obtained from healthy individuals without significant patient risk. When testing an analyte in a patient sample, the developer should be familiar with the most common patient specimens and the optimal maximum time difference between collection and testing. Although the most common patient specimen types depend on the analyte in question and the patient's clinical condition, they generally include (but are not limited to) whole blood, plasma, serum, urine, tissue, nasal swab, and CSF. For many measurands, abundant prior research and literature is available to optimize testing for its intended purposes. The literature supports using particular types of patient specimens, measurand-specific reference intervals, and time intervals for specimen collection. Therefore, when available, patient specimens are the preferred sample type for human clinical studies.

3.2 Surrogate Sample

This guideline establishes a definition of surrogate sample: material or combination of materials used as a substitute for body fluid or tissue taken for examination from a single human subject to study the characteristic of interest. Physical separation, collection into a medical container, multiple collections from the same venipuncture draw, formalin fixation, or paraffin embedding **do not** confer surrogate sample designation. Surrogate analyte, surrogate matrix components, or a combination of surrogate analyte and matrix can be used in surrogate samples. Surrogate samples include but are not limited to:

- Pooled patient samples of biological origin
- Materials supplemented (eg, spiked) with an analyte of interest
- Material created to have properties similar to or representative of the body fluid or tissue of interest
- Material composed of a combination of an analyte that simulates the analyte of interest and a matrix created to have properties similar to or representative of the body fluid or tissue or of the patient or subject
- More-complex combinations of fabricated analyte and matrix

Some proficiency testing materials may be constructed so as to meet the definition of a surrogate sample. Table 2 describes surrogate sample categories.

Table 2. Surrogate Sample Categories⁴

Category	Description
Supplemented	Individual matrix spiked with target analyte
Pooled	Combined or diluted individual patient samples that may or may not be spiked with target analyte
Simulated	Altered materials of biological origin or artificial materials created to have properties similar to or representative of body fluid, tissue, or their components and that may or may not be spiked with target analyte

This page is intentionally left blank.

Chapter 4

Surrogate Sample Hierarchical Approach

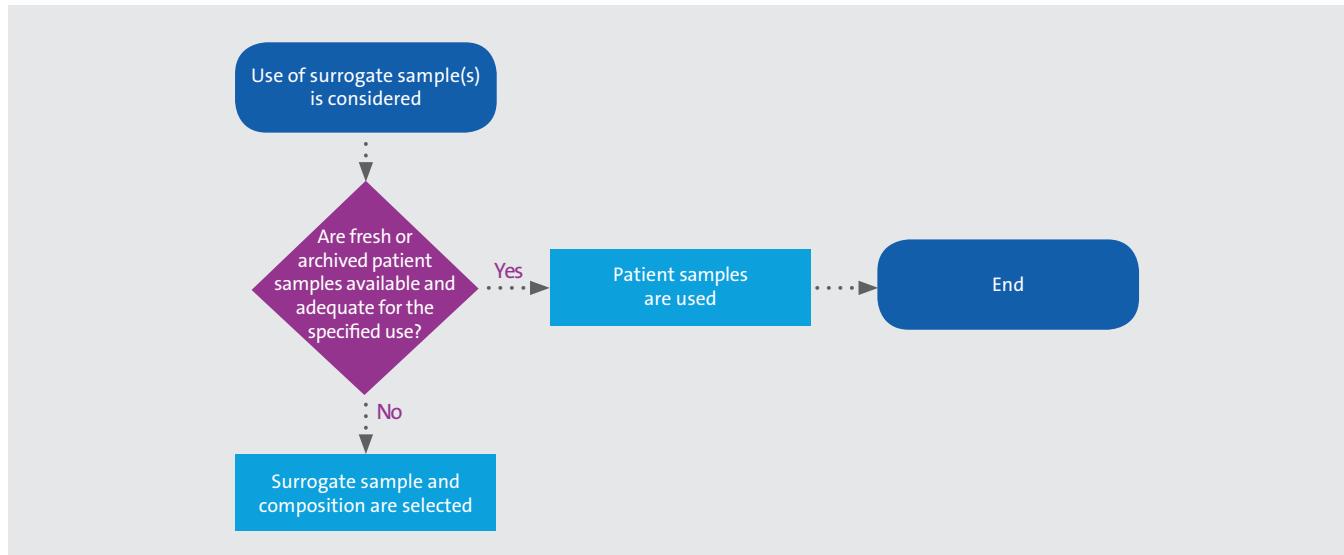
This chapter includes information on:

- Decision to use surrogate samples
- Surrogate sample hierarchy
- Surrogate sample composition

4 Surrogate Sample Hierarchical Approach

4.1 Decision to Use Surrogate Samples

Laboratories and commercial manufacturers (collectively “developers”) prefer to use freshly collected, archived, or frozen patient samples for performance evaluations, some validations or verifications, and regulatory submissions. However, patient samples that are used to test the performance of an assay may be unavailable for several reasons. Figure 2 illustrates the process for deciding to use surrogate samples.



^a Four basic symbols are used in this process flow chart: oval (signifies the beginning or end of a process), arrow (connects process activities), box (designates process activities), diamond (includes a question with alternative “Yes” and “No” responses).

Figure 2. Process for Deciding to Use a Surrogate Sample^a

Depending on the test, the sample may be an unmodified specimen or a specimen that has been processed before examination (eg, use of chemical additives, extraction, centrifugation, or other physical separation methods), which are collectively known as “patient samples.” In general, it is preferable to use surrogate samples to supplement testing conducted with patient samples, rather than relying solely on surrogate samples.

4.2 Surrogate Sample Hierarchy

After deciding to use surrogate samples, developers can apply a hierarchical approach to determine the appropriate surrogate sample type and combination for the designated use. The hierarchy shown in Table 3 minimizes the deviation between surrogate analyte and/or matrix and the patient samples used in the test. Table 3 depicts the possible combinations for surrogate samples and places them in a hierarchy, with pooled patient samples or supplemented individual patient samples at the top. This hierarchy serves as the starting point for composing surrogate samples and can be used, in conjunction with general principles and performance study-specific objectives, to select the best surrogate sample for the planned use or study. The developer should document the decisions and the rationale in a surrogate sample plan, as discussed in Chapter 5. For ease of use, samples in Table 3 have been assigned alphanumerical characters. Appendix A provides surrogate sample descriptions and examples.

Table 3. Surrogate Sample Hierarchy⁴ (Modified from the Medical Device Innovation Consortium (MDIC) Clinical Diagnostics Surrogate Sample Use Working Group. *Surrogate Sample Framework: A Report of the Surrogate Sample Working Group of the Medical Device Innovation Consortium (MDIC)*. Medical Device Innovation Consortium; 2017, with permission.)

			Matrix					
			Biological					
			Unaltered		Altered		Artificial	
Analyte	Unspiked		Patient sample (A1)	Pooled (A2)	Simulated matrix (C1)	Simulated matrix (C2)	Simulated matrix (G)	
	Spiked	Biological	Supplemented (B1)	Pooled (B2)	Simulated matrix (D1)	Simulated matrix (D2)	Simulated matrix (H)	
		Artificial	Simulated analyte (E1)	Simulated analyte (E2)	Simulated analyte/matrix (F1)	Simulated analyte/matrix (F2)	Simulated analyte/matrix (I)	

4.3 Surrogate Sample Composition

Before constructing surrogate samples, the developer needs to choose an appropriate combination of analyte and matrix. See Appendix B for examples of surrogate matrix compositions.

4.3.1 Analyte Options and Considerations

It is preferable to use biologically sourced analytes to create surrogate samples. Of the biologically sourced options, analyte derived from an individual patient sample is generally preferred over analyte derived from pooled patient samples. If individual or pooled patient samples are not available, other biological analytes are used. If biological analytes are unavailable, analytes not derived from samples (ie, artificial analytes with properties similar to the target analyte's) are appropriate. Artificial analytes should mimic the analyte found in a patient sample. Moreover, the developer should consider characteristics such as stability and analyte detection across the AMI. For analytes that are chemical molecules, the approach for biological analytes can be applied.

Additionally, when selecting the analyte source, the developer should keep in mind any particular study design specifications, such as traceability to a standard. When appropriate, standard reference materials, which may or may not be derived from patient samples, can be used as the analyte source. A surrogate analyte source created by pooling or purification typically has the lowest risk of introducing additional factors that may influence test results, such as interferents.

It is common to use a surrogate measurand to characterize endogenous compounds. Strategies include using both synthetic analyte (frequently in an artificial matrix) and a synthetic analog of an analyte (commonly spiked into pooled patient samples). Samples with surrogate analytes are frequently used to prepare samples with known concentrations. These samples may be useful for validation experiments (eg, spike-and-recovery) and method comparison. They also may be used as additional QC samples. Studies using a synthetic analog of an analyte (eg, steroids) have become common.

4.3.2 Matrix Options and Considerations

Matrixes derived from patient samples or unaltered biological matrixes are the media in which the analyte exists and include body fluid or tissue taken for examination. Individual or pooled patient samples are preferred over other sources of patient sample-derived matrixes. If patient samples or patient sample-derived matrixes are unavailable or inadequate, the next best option is altered biological sample matrixes, followed by artificial sample matrixes.

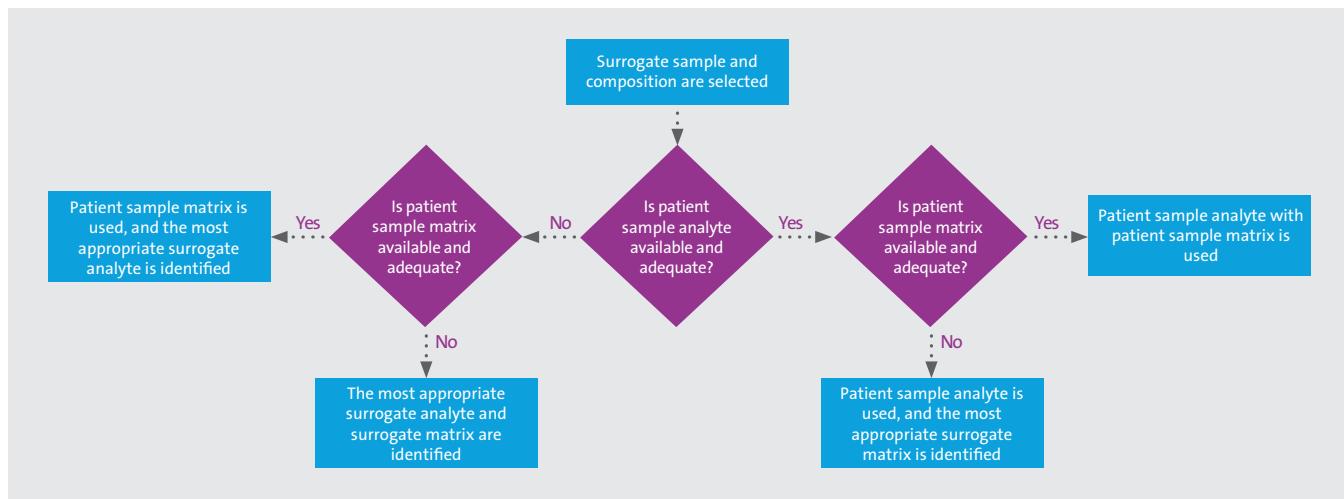
A surrogate matrix created by pooling patient samples typically mitigates the introduction of additional chemical or biological factors that may influence test results, such as interferents. When pooling or diluting, the developer should use the same matrix, if possible, to minimize any matrix effect. For example, a serum sample should be pooled with another serum sample, while an EDTA plasma sample(s) should be pooled with another EDTA plasma sample. For dilution, if the same matrix is not available, the developer should consider adjusting the dilution volume ratio to lower the matrix effect (ie, perturbation of diluent to the sample matrix composition) as much as possible.

Altered biological matrixes have been modified or changed through the application of additional processes not routinely conducted as part of the normal patient specimen collection process. An example of an additional process is treatment of matrixes to remove the analyte or to lower its concentration to an acceptable level, such as creation of a stripped matrix by exposing the patient specimen to activated charcoal.

Artificial sample matrixes are created or manufactured to emulate the patient sample matrix. When preparing an artificial matrix, the developer should consider stability, matrix effects, and measurand detection across the AMI (see Subchapter 7.4 and Appendix B). In some cases, the matrix effect between patient samples and surrogate matrix should be compared in order to predict how well results will correlate. Several methods can be used as needed to determine how well a surrogate matrix will correlate to patient samples, including dilutional linearity, spike-and-recovery, and standard addition (see Additional Resources).^{5,12} These methods can help the developer decide which type of surrogate matrix preparation is most appropriate (ie, dilution vs stripping vs artificial preparation).

4.3.3 Selection of Surrogate Sample Composition

When determining surrogate sample composition, the developer should weigh several factors. Initially, either the analyte or the matrix should be prioritized based on what is needed to emulate a patient sample. Ultimately, the amounts of analyte and surrogate sample matrix should be balanced to reflect their relative importance in mimicking the patient sample, which depends partly on the test type (eg, molecular, immunoassay). Figure 3 illustrates the process for determining surrogate sample composition.



^a Three basic symbols are used in this process flow chart: arrow (connects process activities), box (designates process activities), diamond (includes a question with alternative “Yes” and “No” responses).

Figure 3. Process for Selecting the Composition of a Surrogate Sample^a

Chapter 7 describes additional factors used to select an appropriate surrogate sample, and Chapter 5 provides guidance on developing a surrogate sample plan that uses the principles and strategies described in international standards.¹³⁻¹⁵ The tools described in CLSI document EP23,¹⁶ which uses risk management principles to establish the laboratory's QC strategy, can help the developer decide which surrogate sample formulation is acceptable. Developers should assess factors critical to the study design, including whether the preferred analyte and matrix are available and adequate for use. Additional factors to consider include preparation techniques and storage and handling conditions. The developer should document the selection and suitability of the surrogate sample and the scientific rationale for using it. This rationale should include an evaluation of whether the surrogate sample could interfere with or otherwise influence test performance in a way that patient samples do not. For example, surrogate sample design should minimize effects of sample processing on the test method while mimicking the biological variability of individual patient samples (when possible). Finally, before using surrogate samples, developers should consider the goals and needs of the performance studies. Each performance study should be evaluated independently to determine whether use of surrogate samples is appropriate. Each study is different, and thus one type of surrogate sample may not be suitable for all studies.

This page is intentionally left blank.

Chapter 5

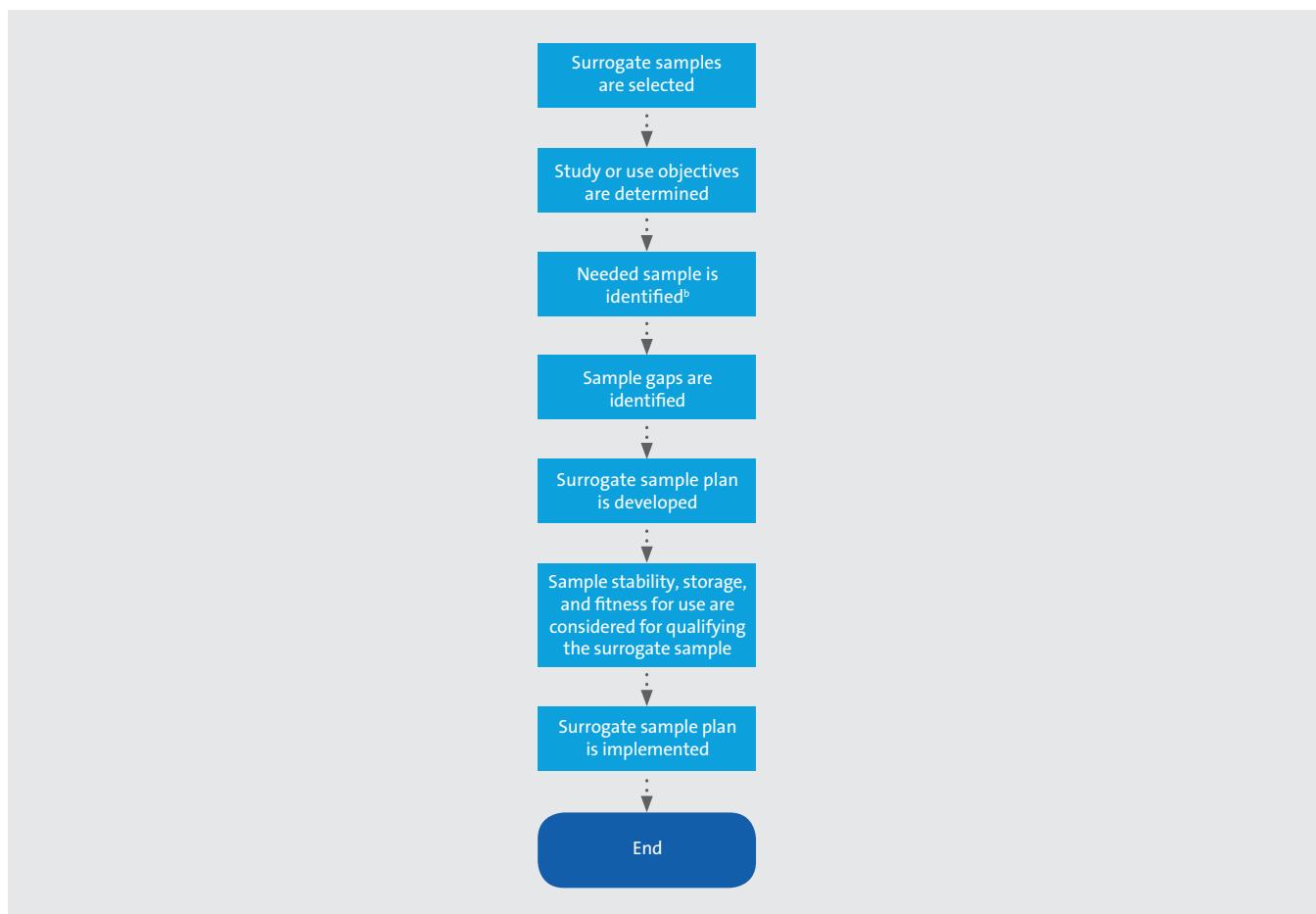
Surrogate Sample Plan

This chapter includes information on:

- Determining study objectives
- Assessing patient sample availability
- Developing a surrogate sample plan

5 Surrogate Sample Plan

Before using surrogate samples, developers should first consider the goals and objectives of each study or use and then determine whether and what type of surrogate samples are appropriate. A surrogate sample type that is suitable for one performance study or use may not be suitable for another study or use. Using the risk management principles described in international standards¹³⁻¹⁵ and in CLSI documents EP18¹⁷ and EP23,¹⁶ the developer can create a surrogate sample plan based on performance study objectives, patient sample characteristics, the principles described in this guideline, and the surrogate sample hierarchy scheme. The plan is used to select the appropriate type and quantity of surrogate samples. It also serves as documentation of the scientific rationale and decision-making process. Figure 4 illustrates the process for developing a surrogate sample plan.



^a Three basic symbols are used in this process flow chart: oval (signifies the beginning or end of a process), arrow (connects process activities), box (designates process activities).

^b The analyte, matrix, volume, and concentration should mimic those of the patient sample.

Figure 4. Process for Developing a Surrogate Sample Plan^a

The relative importance of using the preferred analyte and/or matrix depends on the study type. For some studies, such as method comparison, analyte and matrix may be equally important. For other studies, one may be more important than the other. For a between-assay carryover study, high analyte concentration is essential, but it is also important that the surrogate matrix have appropriate physical characteristics (eg, viscosity) to ensure that the surrogate sample simulates the handling and pipetting of a patient sample. For analytical specificity studies of endogenous analytes (ie, those always present in the patient sample), the developer should prioritize using a surrogate matrix that does not have the targeted analyte (either an altered sample matrix or a patient sample stripped of the analyte). Conversely, interfering substance and cross-reactivity studies may necessitate surrogate analytes to test combinations of specific ratios or concentrations.

Developers should assess the effect of surrogate samples on test performance and determine whether using them could introduce bias. Using surrogate samples as a complete replacement for patient samples across all study types may result in misleading conclusions regarding test performance. Generally, surrogate samples should be used to augment testing performed with patient samples, as opposed to completely replacing patient samples. Additionally, in some situations, it may be preferable to use an alternative approach rather than a surrogate sample. For example, a separate small-scale study with patient samples can be supplemented with a larger-scale study that includes surrogate samples.

5.1 Identifying Needed Samples

After designing the study, the developer needs to evaluate the quantity of available patient samples, as well as the sample volume needed to conduct the study and any repeat tests that may be needed. This evaluation includes an assessment of the availability of both the negative and positive patient samples that would be needed to meet the study design criteria. For endogenous analytes, the developer should evaluate the availability of suitable negative patient samples. For analytes with various genotypes or serotypes, the availability of patient samples containing the various types should be assessed. For multiplexed tests that interrogate mutations of interest, the developer should ensure that samples containing all the common mutations and relevant combinations are available. Additionally, developers should ensure that sufficient sample volume is available when it is likely to be limited. Factors that may make obtaining adequate sample volume difficult include:

- Patient sample type (eg, CSF, tears)
- Patient population (eg, pediatric patients)
- Number of replicates necessary to complete the study (eg, reproducibility study)

For qualitative tests, the developer should determine the number of negative and positive patient samples necessary to characterize test performance around the lower limit of detection (LLoD) or cutoff (refer to CLSI documents EP12¹⁸ and MM03¹⁹). For quantitative tests, developers should calculate the number of patient samples needed to cover the AMI (refer to CLSI document EP06²⁰ for an example of a linearity study), including around the LLoD and lower limit of quantitation (LLoQ) (refer to CLSI document EP17²¹) and around the MDLs (refer to CLSI documents EP09²² and MM06²³). Paired samples (ie, different specimen types [eg, serum, EDTA plasma, heparin plasma] taken from the same patient) may be needed for sample matrix comparisons of similar sample types. After thoroughly assessing patient sample availability and study needs, the developer should determine the gaps in patient samples, for which surrogate samples will be needed to complete the study.

5.2 Developing a Surrogate Sample Plan

A surrogate sample plan documents the selection of surrogate samples suitable for the study's needs. Although the use of surrogate samples is well established for some technologies, analytes, and studies, it has been less frequent in other areas. Using risk management principles, developers should document the scientific rationale, which may include experimental data, literature references, previous studies, and/or other information, that supports the use of the identified surrogate sample type (see CLSI documents EP18¹⁷ and EP23¹⁶).¹³⁻¹⁵ The surrogate sample plan can be tailored to the situation. For example, when there is an existing body of knowledge, sufficient guidance, or a recognized established surrogate sample, a simplified plan with fewer elements may be appropriate. Developers of the surrogate sample plan should:

- Describe the study type or use (eg, performance study, release testing).
- Describe the study or use objectives.
- Describe the samples needed and the gaps in sample availability.
- Document the rationale for using surrogate samples (eg, scientific demands of study type, prevalence of patient samples).
- Describe the hierarchical approach used to select the surrogate sample type (see Table 3 and Chapter 8).
- Describe the patient sample characteristics that are most relevant for the study.
- Describe the degree to which a surrogate sample is changed from a patient sample.
- Evaluate how well the surrogate sample mimics the biological variability of individual patient samples.
- Consider the effect of sample processing (eg, extraction methods), if applicable, and determine an approach to preparing surrogate samples that mimic patient samples when processed.
- Evaluate the stability of the surrogate samples (see Subchapter 6.4).
- Determine the aspects of analyte and matrix combinations that should remain unchanged and the aspects that may be altered, as well as the extent of alteration. The relative importance of using biological analyte vs biological matrix depends on the type of test (eg, molecular, immunoassay), patient sample, and performance study. The developer should justify the surrogate analyte and matrix choices and assess and document the risks involved with use of such substitutes.
- Qualify surrogate samples for use to ensure that they meet the study needs. This process may include confirmation of analyte-free concentration or stability (see Subchapter 7.1).
- In the absence of a supporting scientific rationale, establish comparable performance of surrogate samples and patient samples (see Subchapter 7.2 and CLSI document EP35²⁴). The concept of commutability may be applicable to demonstrating comparable performance of patient and surrogate samples between or among quantitative test methods (see CLSI document EP14²⁵). If analyte-free matrix is not available, a matrix equivalence study is generally used to support the use of non-analyte-free matrix. Furthermore, surrogate matrix may not be appropriate in all cases.
- Consult existing CLSI guidelines regarding the use of surrogate samples specific to the study type, such as CLSI documents EP05²⁶ (precision), EP06²⁰ (linearity), EP07²⁷ (specificity), EP14²⁵ (commutability), EP17²¹ (analytical capability), EP34²⁸ (extended measuring interval), EP35²⁴ (sample suitability), and C37²⁹ (frozen human serum pools).
- Use the hierarchical approach to select an appropriate surrogate sample (see Table 3).

Chapter 6

Technical Preparation Techniques

This chapter includes information on:

- Analyte and matrix determination
- Surrogate sample preparation
- Stability and storage

6 Technical Preparation Techniques

6.1 Overview

To reduce process errors, surrogate samples should be prepared by qualified developers. Equipment should be qualified with recognized laboratory standards to ensure suitability for use. The developer should follow standard laboratory practices and procedures to minimize the risk of producing poor-quality surrogate samples. Operator training should review good laboratory practices that minimize the risk of cross-contamination and sample loss, including the use of personal protective equipment, proper pipetting and measurement techniques, proper mixing condition(s), and a clean workspace. Personnel should be trained and/or certified in the processes they will perform. Using disposable labware or glassware (if available and appropriate) is a good option for limiting cross-contamination. The developer should follow standard cleaning and storage procedures for labware.

Preparation techniques for surrogate samples vary (refer to CLSI document EP06²⁰ and international standards³⁰). The appropriateness of a surrogate sample for its intended purpose depends partly on its physical properties, which should be considered during selection of both analyte and matrix. These physical properties include stability, solubility, pH, viscosity, specific activity, purity, and phase changes (eg, during storage or from saturation of the material in solution). When preparing a surrogate sample, the developer should consider the order of addition of the component parts (eg, gravimetric addition to solvent, solvent to solid addition, volumetric mixing of solvents or suspensions). The developer should also consider mixing technique (eg, inversion, vortex, magnetic stirrer, multiple pipetting), speed, temperatures of materials, process parameters (eg, mix and centrifugation times), filtration media, and storage temperature of components and of the final surrogate composition. For proteinaceous materials, mixing technique and speed are especially important.

Chapter 5 discusses identifying study objectives, which help define process steps related to the composition of the surrogate sample. For example, if a surrogate sample will be used as an internal reference point for ongoing manufacturing capability or product release, the developer may need to establish acceptable ranges. If the surrogate sample will be used to evaluate precision and comparison studies, adequate storage and handling properties, including the sample's robustness to freeze-thaw cycles, are important considerations.

6.2 Determining the Analyte

In principle, as discussed in Subchapter 4.3.1, the analyte in a surrogate sample should be as close as possible, with regard to characteristics and properties, to the analyte present in patient samples. In many cases, the chemical composition of the analyte to be used is identical to that found in patient samples. For example, NaCl, lactic acid, ascorbic acid, and small-molecule drug preparations are well-characterized chemical compounds. When the analyte's chemical structure is known, using high-purity material can increase confidence in the adequacy of preparation. In other cases (eg, need for increased stability, rarity of analyte), the developer may use recombinant proteins, although they might differ in structure (eg, in secondary or tertiary configuration) or lack post-translational modifications. Similarly, apoenzymes (ie, inactive enzymes) may be used in place of the holoenzyme. For nucleic acid assays, smaller segments of DNA or RNA or oligonucleotides may be considered adequate surrogate analytes, depending on the study objective(s).

Different analyte types (eg, recombinant protein vs naturally occurring protein) can result in varying levels of test method bias. For example, with ELISA, a recombinant protein may be appropriate for a reproducibility study but not for detection. The developer should consider different surrogate analyte types in combination with surrogate matrixes. The needed analyte concentration should also be a consideration. If a high analyte concentration is

needed, the developer can use microorganisms grown in culture or synthetic nucleic acids or similar molecules that can produce high analyte concentrations.

Considerations for specific performance studies include:

- **Linearity:** Some studies, such as linearity, may necessitate creating surrogate samples composed of analyte dilutions distributed across the AMI (refer to CLSI document EP06²⁰). In such cases, it may be appropriate to create surrogate panels spanning the AMI by spiking varying amounts of a patient sample containing a high analyte concentration into a negative patient sample.
- **Detection capability:** The developer may assess the availability of reference materials for the analyte or consider using biospecimen or organism banks (see Additional Resources). Reference materials can be used to create test panels for studies such as detection capability.
- **Analytical specificity:** To assess the specificity of tests for endogenous analytes, the developer can construct suitable blank surrogate samples by stripping the analyte from native patient specimens (eg, through precipitation by an antibody, enzymatic degradation, or adsorption to charcoal), although such sample manipulation may risk altering other constituents (eg, binding proteins) important to the test method (refer to CLSI document EP17²¹).

Using gravimetric addition or volumetric dilutions and admixtures enables the developer to define absolute or relative analyte concentrations. If the method includes an extraction step, the extraction medium may serve as the surrogate sample, depending on the study's purpose(s). For example, drug analysis by chromatographic or mass spectrometric techniques often requires cleaned-up samples to establish the instrument's detection range. In addition, for some microbiological samples, the analyte (eg, antigen or antibody) is extracted before being applied to the test used for detection or quantitation. In these cases, the developer should consider handling, storage, and use with regard to the study's purpose(s).

Moreover, the source of analyte used to prepare surrogate samples is often a high-concentration stock solution and/or a solution with a nonphysiological solvent (ie, not similar to the sample matrix). When such a solution is used to prepare the surrogate sample, the added volume of stock should be as low as possible, so as to cause minimal disruption to the surrogate matrix composition used for testing. The developer should screen the surrogate matrix for the analyte before spiking. Furthermore, surrogate samples prepared for a highly sensitive assay that can measure or detect very low amounts of analyte may need very low ratios of analyte to diluent. Correct preparation of such samples may depend highly on volume handling techniques (eg, volume transfer). It is strongly recommended that developers be qualified and use certified equipment to prepare these samples. The same is true for mass spectrometry and certain molecular techniques.

When the analyte is labile, the developer may need to use fresh preparations of the surrogate sample to ensure that the sample is adequate for the study's purpose(s). In such cases, special storage conditions may be needed (eg, temperatures $\leq -20^{\circ}\text{C}$ in a freezer, use of liquid nitrogen to lower temperatures). Examples of labile analytes are protease-sensitive proteins and hemoglobin, which converts to ferrihemoglobin unless specially handled and stored.

The source of the surrogate analyte may have the potential to introduce interferences (eg, cross-reacting substances, inhibitors). The developer should always consider the appropriate analyte source for the experiment's or evaluation's specific purpose(s).

6.3 Selecting the Matrix

When selecting the surrogate matrix, the developer should consider the properties and characteristics of the patient sample that are most relevant to the study type. Developers should also consider aspects of the surrogate sample that would interfere with or otherwise influence test performance in a way that is inconsistent with patient samples. For example, if matrix effects (ie, biases caused by unidentified materials, components, or characteristics of the matrix) are rare, pooling sources is generally acceptable. For many studies, patient samples are commonly pooled for matrixes such as serum, plasma, and urine. However, if the matrix and analyte interact negatively or if the patient sample matrix varies, it is generally advisable to limit the number of individual samples and/or matrixes, so that the pool contains the fewest individual patient samples necessary to achieve the required volume. For example, nasal swabs and fecal material tend to vary more than serum. The developer should consider the amount of variability in the matrix. Especially for highly variable matrixes, minimizing the number of samples in the pool helps preserve the biological composition of the matrix. Matrix-induced biases are often more efficiently evaluated in comparison studies than in studies that analyze a single matrix in isolation (eg, through chromatography). Such information helps identify the inherent advantages and disadvantages of surrogate samples. The developer should also consider matrix components (eg, proteases, lipases, enzyme inhibitors) that have the potential to degrade the analyte or other matrix components. Refer to CLSI document C49³¹ for guidance on matrix effects.

When using an artificial matrix, the developer should describe the characteristics of the patient sample that are represented or mimicked by the surrogate sample. Developers should also consider other factors (eg, viscosity, endogenous cellular components) so that differences between patient and surrogate samples are minimized. The developer should establish the equivalence of the patient and surrogate matrixes (refer to CLSI documents EP14²⁵ and EP35²⁴). When the surrogate matrix differs significantly from the patient sample matrix, the developer should document the specific performance characteristics to be assessed, along with justification for using the chosen surrogate matrix. For example, use of organic solvents may be the only way to solubilize a particular drug or hormone. If solubility may affect the assay's reliability, a soluble polymer can be used to increase a sample's viscosity.

6.4 Surrogate Sample Stability

It is important for the developer to consider the stability of surrogate samples (refer to CLSI document EP25³² [which applies to reagent stability but may be useful to EP39 users] and international standards³³). Often, surrogate samples need to be stable under storage conditions and for time periods that differ from those required for patient samples in routine clinical use. Using accepted methodologies, the developer should design surrogate sample stability studies to validate stability requirements, including those for storage stability at varying temperatures and freeze and thaw stability.

The surrogate sample stability requirements depend on the study type. For example, linearity studies generally require samples that are stable for a short time, while precision panels (often used for 20 or more days) may require samples with longer stability. The characteristics of the analyte or matrix sometimes make surrogate panels unstable. When materials other than patient samples (eg, artificial analyte or biological analyte grown in culture) are used as an analyte source, the analyte may degrade when spiked into patient sample matrix.

The developer should establish the stability of the patient sample matrix that has been altered for use as negative matrix for the purpose of detection studies in advance so that result variability is not introduced through matrix degradation. For example, stripping patient samples of the analyte through size exclusion or charcoal treatment may create instability. Analyte stripping may also change the matrix properties for low-molecular weight analytes, making the use of stripped matrix unreliable (eg, for assessing the LLoD and/or LLoQ).

For large sample pools, the developer should establish the stability of the surrogate matrix and analyte so that variability in the characteristic of interest is not introduced through sample degradation. Acceptable variation ranges for surrogate sample stability can be identified by assessing analyte stability in a known matrix. Before examination, surrogate samples should be processed with the same preexamination steps as patient specimens, so that supplemented analyte is processed equivalently to the analyte in patient specimens.

6.5 Process Steps

The developer should detail process steps such as mixing techniques in the preparation documentation. For example, the document should specify whether only (small-scale) inversion is appropriate or whether low-speed mechanical mixing is needed to avoid frothing, which could induce analyte or matrix degradation. Depending on the surrogate sample's intended use, additional process steps may be needed for opaque or viscous samples. Testing the concentration of the final solution (if possible) can help ensure that the target concentration was achieved. The developer may need to evaluate and test critical parameters such as pH, conductivity, and density.

6.6 Storage

If the surrogate sample will be used for long time-interval assessments, the developer should determine the number of freeze-thaw cycles the sample will undergo, in order to ensure continued adequacy of the analyte (including solvents [for stock solutions], if applicable) and of the matrix. Additionally, the developer should consider the storage container used for the surrogate sample, which has the potential to affect sample integrity and performance. For example, if surrogate samples will be frozen for long-term storage, developers may use cryogenic vials or tubes to maintain sample integrity. The developer should indicate the type of glass or plastic composition selected, because many containers have highly charged surfaces that can bind the analyte of interest or other matrix components, damaging the surrogate sample's integrity for its intended purpose(s). A carrier-blocker molecule (eg, serum albumin) is routinely added to synthetic matrix to prevent analyte adsorption to plastic containers. Lyophilizing the material is an option for long-term storage, but this practice can result in variable analyte concentration, owing to the process of reconstitution. Instead of low-density polyethylene or glass storage containers, the developer could use containers made from high-density polyethylene, which is strong and thus resistant to shattering during the freeze-thaw process. The use of blackout or amber tubes may be appropriate to maintain the integrity of light-sensitive analytes.

This page is intentionally left blank.

Chapter 7

Special Considerations

This chapter includes information on:

- Qualifying the surrogate sample
- Demonstrating comparability
- Using samples for more than one study
- Preparing blank samples
- Creating a single signature score from multiple genes
- Rare subtypes
- Surrogate samples for molecular assays

7 Special Considerations

7.1 Qualifying the Surrogate Sample

The scientific rationale for using a particular surrogate sample type, including experimental data, literature references, previous studies, and/or other information, can be used to demonstrate comparability of the surrogate sample to patient samples. In the absence of a supporting scientific rationale, the developer may conduct a study to demonstrate comparability between patient and surrogate samples. The ability to conduct such a study depends on the availability of patient samples with the concentrations to be tested. Comparability can only be demonstrated at available concentrations of patient samples.

Developers should ensure that surrogate samples meet the study specifications (ie, targets, recovery ranges, sample characteristics, and performance) before use. After a surrogate sample has been prepared, qualification can confirm its suitability for use and/or its comparability to a patient sample. A surrogate sample type that is suitable for one performance study may not be suitable for another. The total preparation volume should be sufficient for qualification testing as well as the planned studies.

Typical qualification studies and/or assessments are:

- Analyte concentration (ie, target concentration or concentration intervals)
- Analyte and surrogate matrix stability (ie, normal storage stability, freeze-thaw stability, and room temperature stability)
- Sample comparability (ie, performance of surrogate samples vs patient samples or well-recognized standards [see Additional Resources])
- Sample physical conditions (eg, volume, pH, viscosity)

Approaches for qualifying samples are:

- Value assignment test
 - The developer identifies the target value and acceptable range before beginning the preparation process. Processing and final testing of surrogate samples should include enough replicates and runs to achieve adequate statistical power to confirm the target value and/or range (when a reference test method exists). Tests can be performed on multiple platforms, systems, or test lots, based on each study's requirement(s).
- Stability evaluation (see Subchapter 6.4)
- Comparability study (ie, performance of surrogate vs patient samples) (see Subchapter 7.2)
- Analytical testing (eg, gas chromatography–mass spectrometry, high-performance liquid chromatography, certificate of analysis [for purchased samples traceable to a quality standard])
- Commutability studies
 - The developer can evaluate commutability when qualifying surrogate samples for certain quantitative tests. For example, commutability studies may be useful during development of a universally available stable reference material. Commutability can also be assessed for stable sample panels directly traceable to clinical conditions, created for the purpose of demonstrating a diagnostic test's clinical validity (see CLSI documents EP14²⁵ and EP30³⁴).³⁵⁻³⁷

For the surrogate analyte approach, qualification may include demonstration of similar performance in the patient and surrogate matrixes after completion of examination steps (eg, extraction recovery). Similarly, for the surrogate matrix approach, qualification may include demonstration of similar extraction recovery in the patient and surrogate matrixes.

7.2 Demonstrating Comparability

The type of study conducted to demonstrate sample comparability depends on the test type (eg, quantitative vs qualitative) and the surrogate sample's intended use. For quantitative tests, method comparison or commutability study designs may be appropriate in some cases, while in other situations, sensitivity, repeatability, and linearity studies may be more suitable. For qualitative tests, sensitivity (near the LLoD and/or cutoff) and repeatability studies may be appropriate. Assessment of comparability follows the general principles of method comparison studies. Subchapters 7.2.1 and 7.2.2 provide examples of study designs for demonstrating sample comparability for quantitative and qualitative tests, respectively.

7.2.1 Quantitative Tests

Method comparison studies for quantitative tests typically use surrogate samples to supplement the data for abnormally high test values, which become progressively rarer as the analyte concentration increases. However, lower measurement variability for the surrogate samples vs the patient samples could result in an overly optimistic confidence interval for the regression slope. In order to more easily assess any differences between patient and surrogate samples, patient samples should overlap with surrogate samples across the intended AMI, as depicted in Figure 5.

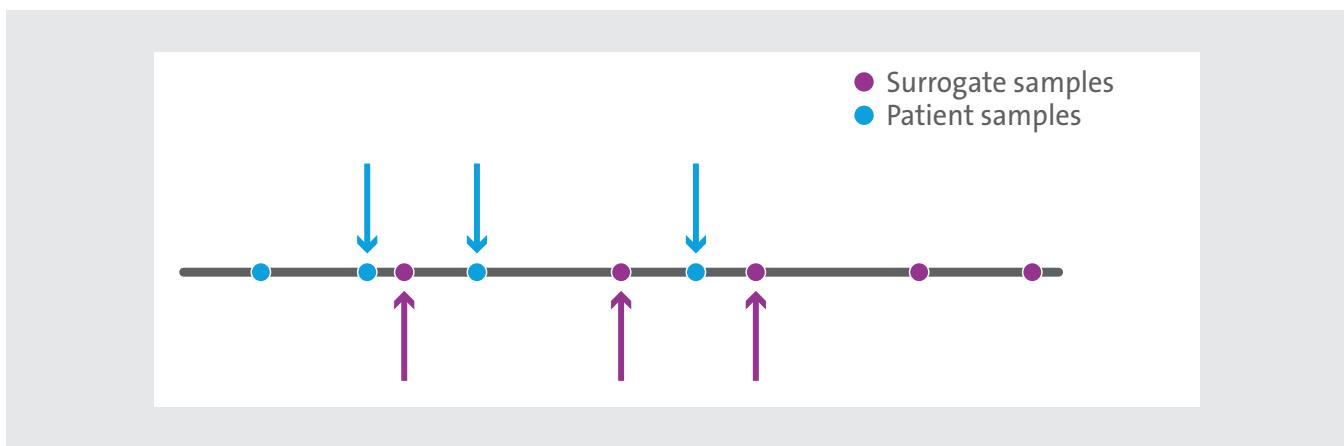


Figure 5. Comparison of Patient and Surrogate Samples With Overlapping Concentrations. The horizontal axis indicates concentration. Arrows highlight the overlapping concentrations between patient and surrogate samples.

The recommended statistical analysis is linear regression that is appropriate for the analytical performance and clinical indications of the candidate measurand. Examples include simple linear, weighted Deming, and Passing-Bablok. The developer should perform analysis of residuals from the corresponding regression lines of patient and surrogate samples, calculated separately. For example, if the surrogate samples fall within the 95% prediction interval of the linear regression line for only the patient samples, the surrogate samples are comparable. Analyte concentrations should span the claimed AMI, including at relevant medical decision points. When the surrogate samples do not satisfy these criteria, the developer should assess the clinical significance

data, along with the surrogate sample's intended use and the likely clinical effect of not satisfying the sample criteria. In addition, it may be useful to prepare a contingency table (eg, 2 × 2) to calculate the percent agreement and confidence intervals for each MDL (refer to CLSI document EP12¹⁸). For more information on changes in test interpretations, refer to CLSI document MM17.³⁸

7.2.2 Qualitative Tests

The techniques used to assess comparability of patient and surrogate samples for method comparisons of quantitative tests are not applicable to qualitative assays. As discussed in Subchapter 8.6.5, sample comparability in qualitative assays can be assessed by calculating the positive percent agreement (PPA) and negative percent agreement (NPA) separately for the patient samples and for the surrogate samples. It is important that the percentage of samples with concentrations close to the cutoff (eg, $\approx 30\%$) be similar in patient and surrogate samples. The distribution of analyte concentrations for the remainder of patient vs surrogate samples should also be similar. Because it is often impossible to assess the signal intensity from qualitative tests, the prepared surrogate samples should span the AMI throughout the expected distribution of concentrations for the intended-use population. Ideally, a significant percentage of the surrogate samples should have concentrations near the LLoD or cutoff (refer to CLSI documents EP12¹⁸ and MM03¹⁹), although this objective is often difficult to achieve. If the PPA and NPA are similar for patient and surrogate samples, the surrogate samples are comparable. If the surrogate samples represent a small fraction of the total dataset, supplementing the method comparison study with the surrogate samples is unlikely to introduce a bias. The surrogate sample dataset may have a statistical effect on calculations such as the PPA. Thus, when evaluating the total sample set (ie, patient and surrogate samples), the developer should evaluate the surrogate sample dataset for a clinical effect.

Frequently, validation studies for qualitative tests include samples with concentrations at approximately two to three times the LLoD. These low-concentration samples may be necessary to ensure that the test method can detect very low analyte concentrations under different testing conditions. If individual patient samples with low concentrations are difficult to source, the developer should consider using surrogate samples. For study design considerations, refer to CLSI document EP12.¹⁸

7.3 Using Surrogate Samples for More Than One Study

Preparing and qualifying surrogate samples can be difficult and time consuming. Therefore, being able to use a surrogate sample in more than one study is helpful. The developer should determine all intended uses of the surrogate sample before designing characterization studies. For example, a negative surrogate sample can be used as the low pool for sensitivity and linearity studies, in which the absence of the analyte is desired. Characterization studies ensure that the surrogate sample (with zero or low analyte concentration) can be used for sensitivity and linearity studies, which require low analyte concentrations. Although surrogate samples with low analyte concentrations are not useful for a limit of blank (LoB) study, they can be used as a source of matrix for low-concentration surrogate samples (eg, for interference and precision studies). Similarly, for high-concentration precision and linearity studies, a matrix with a higher analyte concentration is desirable because it reduces the need for a very high-concentration stock solution for spiking. If an analyte stock is available only in limited volume and concentrations, determining the studies in which it will be used is helpful. In this way, the developer can determine whether the surrogate sample is sufficient for all studies or whether supplemental material (eg, a matrix with a high analyte concentration) is needed.

7.4 Preparing Blank Surrogate Samples

Preparation of blank surrogate samples is often necessary to evaluate measurand performance in various validation studies. In some cases, if the analyte concentration is at least an order of magnitude below the LLoD, the patient sample can be used (see CLSI document EP17²¹). In other situations, even a low amount of analyte can be problematic (eg, polymerase chain reaction). In principle, blank surrogate samples should be similar to the patient samples, differing only by the absence of the analyte. Thus, a blank surrogate sample should have properties (eg, cellular content, viscosity, chemical composition) similar to or representative of the body fluid or tissue used for the test.³⁹

When patient samples are not available (owing to limited volume or other constraints), blank surrogate samples can often be used to establish analytical detection limits, determine the effects of interference, or evaluate linearity (at the lower limit of the measuring interval [LLMI]). For validation studies, a high volume of blank surrogate sample is often needed to create the necessary surrogate samples. The developer may need to pool samples, with or without diluting multiple patient samples, or use other appropriate surrogate samples (whose matrixes closely resemble individual patient sample matrix).

Patient samples that are analyte free or at least an order of magnitude below the concentration of interest for the analyte(s) (see A1 in Table 3) are the preferred blank patient samples. If this sample type cannot be used, the developer should choose the preferred surrogate sample, following the hierarchy shown in Table 3. This order of preference is based on the principle of maintaining individual sample variability to better represent patient sample performance. If the needed volume cannot be obtained from an individual sample, the next best option is to pool as few patient samples as possible, thus preserving the biological composition of the matrix (refer to CLSI document EP17²¹). If a known analyte-free patient sample is difficult or impossible to obtain, the developer may consider using a simulated artificial matrix that closely resembles the patient sample. An artificial matrix mimics the patient sample composition (eg, cellular background, normal site flora) and has protein, electrolyte, and pH properties similar to those of the patient sample.

Based on the same logic, preferred blank surrogate samples can include patient samples from a population that does not typically express the analyte (eg, β -human chorionic gonadotropin [β hCG]-free male serum pools for evaluation of β hCG). A similar example is a patient sample (from the same anatomical source) from a noninfected individual in whom testing has confirmed the absence of all pathogens detected by the assay. The developer can supplement blank patient samples from that individual with the representative pathogens and analyze the samples accordingly.

In some cases, the developer can use protein precipitation, liquid-liquid extraction, or solid-phase extraction to remove the analyte from patient samples, thereby creating a blank surrogate sample. Immunoprecipitation techniques can be used to create a blank surrogate sample for evaluating the performance of an immunoassay or other test. For example, the developer could incubate the microparticle coupled (specifically for the reagent antibodies) with the low-concentration sample and strip the analyte through magnetic attraction. This procedure renders the low-concentration sample supernatant essentially analyte free, meaning that it could serve as a blank surrogate sample. However, stripping a patient sample of an analyte risks changing its matrix composition.

Appendix B provides examples of artificial matrix compositions, which can be used to guide preparation of blank surrogate matrixes for bronchoalveolar lavage, CSF, intestinal fluid, saliva, semen, sputum, sweat, synovial fluid, tears, urine, vaginal fluid, and serum. In some cases, synthetic samples (eg, synthetic urine) can be purchased. These examples provide a reference point for the composition of a specific matrix. For instance, the example artificial sweat matrix described in Appendix B does not represent the only possible formulation (eg, additional formulations of artificial sweat can also be prepared).^{40,41} Similarly, a different artificial matrix may be needed, depending on sample timing (eg, fasted-state vs fed-state simulated intestinal fluid).⁴²

The presence of endogenous elements can be beneficial for evaluating analyte performance in the intended-use population. However, in excess, these substances can cause a matrix effect for specific technologies, which may create bias in the study results. For example, high biotin blood concentrations may affect some immunoassays that rely on avidin-biotin binding. Patient samples should be free of obvious interferents that may generate bias. Nonetheless, surrogate samples should adequately reflect the complexities in the desired sample type, and the matrix should not be artificially simplified.

Analyte-free surrogate samples can be spiked with low analyte concentrations. Considerations for specific situations include:

- **LLoD or LLoQ studies:** LLoD and LLoQ studies are generally performed in unaltered biological sample matrix, because these parameters support other performance studies (eg, interference) (refer to CLSI document EP17²¹).
- **Rare matrixes:** Patient samples may be unavailable or difficult to obtain (eg, tears). Thus, an artificial matrix may be appropriate.
- **Interference studies:** The goal is to determine whether the assay can specifically detect the target analyte in the presence of relevant interferents. Therefore, potential interferents are spiked into individual blank samples to evaluate the effect of the elevated interferent levels on the assay (see CLSI documents EP07²⁷ and EP37⁴³).

The use of artificial matrixes is well established for some technologies, analytes, and studies. In the surrogate sample plan, the developer should document the scientific rationale, which may include experimental data, literature references, previous studies, and/or other information, that supports the use of a particular surrogate sample type, including the artificial matrix and other considerations. When existing data are lacking, the developer may support the use of the artificial matrix by performing a matrix comparison study to evaluate its appropriateness. A matrix comparison study compares analyte performance in natural (ie, patient sample) vs artificial sample matrix. The study design depends on the availability of the natural sample matrix. If sample volumes are sufficient, the study should generally include analyte concentrations above and below the LLoD. When the matrix will be used to dilute positive analytes, testing below the LLoD may not be necessary. The developer should evaluate the differences in performance between the natural and artificial sample matrixes, based on rational assessments and data analysis. The purpose of a matrix comparison study is to support the use of an artificial matrix. In some cases, the study may find that the artificial matrix is not appropriate. The use of nonhuman specimens (eg, horse blood) as an artificial matrix may not be acceptable, because biological differences between human and nonhuman serum may create nonspecific reactions. While Appendix B provides information on the basic chemical composition of artificial matrixes, Table 4 lists considerations for specific surrogate sample types.

Table 4. Example Factors to Consider When a Blank Matrix Is Chosen

Surrogate Sample	Considerations
Respiratory, BAL, sputum, nasopharyngeal	Negative human clinical respiratory specimens can be pooled to create a high volume of uniform sample matrix (eg, negative nasopharyngeal pools prepared from leftover nasopharyngeal swabs, BAL, and sputum). These matrixes are more representative of the patient matrix than an artificial matrix would be, because they contain the normal flora organisms and cells present in those sample types. When preparing an artificial matrix for these sample types, in addition to creating the matrix's chemical composition, the developer may add nontarget normal flora organisms and a cell line to mimic the background organisms found in patient BAL specimens.
Vaginal fluid	Various organisms may be present in vaginal fluid specimens. Because natural vaginal matrix may contain the analytes of interest as part of the normal vaginal flora, the developer may determine the LLoD separately for each analyte in artificial matrix only, as appropriate. However, the LLoD should be confirmed separately in natural matrix and simulated matrix to ensure accuracy.
Cervical tissue for HPV testing	An HPV-negative cell line may serve as a blank surrogate sample for cells not infected by HPV, while LBC samples contrived from HPV-infected cell lines (eg, SiHa, HeLa) are used as the positive surrogate samples. The developer may perform a paired-sample LLoD study, using at least 1 HPV-infected cell line, to demonstrate that pooled negative clinical matrix and simulated background matrix (ie, an HPV-negative cell line in LBC media) produce the same LLoD results. If these 2 samples demonstrate equivalence, the HPV-negative cell line can be used as the background matrix in other validation studies.
Serum	Blank serum can be used to prepare known samples (eg, spike-and-recovery samples, calibrators). For some methodologies, components such as buffer or salt may need to be adjusted or replaced.
Blood culture	The LLoD can be determined by preparing serial dilutions of positive blood cultures and spiking them into negative simulated matrix (eg, blood culture media spiked with human blood).

Abbreviations: BAL, bronchoalveolar lavage; HPV, human papilloma virus; LBC, liquid-based cytology; LLoD, lower limit of detection.

7.5 Creating a Single Signature Score From Multiple Genes

Although this subchapter primarily concerns molecular tests, it can be applied to other multianalyte tests as well. This subchapter focuses on surrogate analytes, as opposed to surrogate matrixes. Analyte measurement for multianalyte molecular tests varies based on numerous factors, including disease stage, health status, genetic variation, and patient age. For example, disease onset for one patient may elicit a response that produces very high concentrations of all analytes. Conversely, for another patient, only a subset of analytes reach the same concentrations. Considerations for design and development of surrogate samples for quantitative multianalyte tests that report a single value include:

- **General guidance for performance studies:** In general, the recommendations in Chapter 8 for using surrogate samples for performance studies apply to multianalyte molecular tests.
- **Score vs analyte:** Some studies are relevant to the score (eg, interference, cross-reactivity). Others are relevant to the analyte and necessitate that each tested analyte be assessed individually (eg, analytical sensitivity, linearity).
- **Same result for different analyte combinations:** For tests that compare a score with a threshold, it is possible for different combinations of analytes to lead to the same score.

Because different combinations of individual analytes can lead to the same overall result, it is recommended that the developer include all realistic combinations of analytes for each desired result. In some cases, developers may need to use surrogate samples to obtain a certain combination of analyte concentrations yielding a particular score that is possible for patient samples. However, multiple combinations of individual analytes in patient samples that lead to the same combined result are preferred. For example, using multiple samples with the same or similar scores is encouraged, because they may confer different levels of precision for that score.

When determining the sample combinations to test, the developer needs to understand how individual variables are considered and combined to determine the final result (eg, whether some variables are weighted more heavily than others). Moreover, for endogenous markers, it is important to understand which values constitute normal results vs disease-related changes. When obtaining several analytes in a pure form is impossible (because the analytes cannot be obtained or synthesized individually), it is acceptable to use a patient sample that contains all the analyte derivatives, rather than creating a surrogate for each individual analyte. When creating all combinations is not practical (eg, there are many analytes), it is recommended that the developer create multiple clinically relevant combinations per result.

In order to define “clinically relevant” combinations, it is important to consider the biology of the intended-use population. For example, in non–small cell lung cancer (NSCLC), it is possible to observe multiple positive fusions. However, it is extremely unlikely that an assay designed to detect nonreceptor tyrosine kinase fusions would identify multiple fusions in the same patient. Therefore, creating such a combination would defeat the purpose of a surrogate sample. However, in some cases, evaluating such a surrogate sample may be desirable (eg, confirming that the analysis logic will not falter if such a sample is encountered).

7.6 Rare Subtypes

Generally, developers use standards and patient samples to establish performance, such as the LLoD, LLoQ, and linearity of a specific measurand, often using the most prevalent type of analyte. It is important to confirm that these parameters are the same for all genotypes, subtypes, and other variants of the analyte. For qualitative assays, the developer can confirm detection limits and linearity by testing surrogate samples created through dilution. Patient samples (or isolates in patient sample matrix or equivalent) are diluted to low-positive concentrations (eg, near the cutoff or two to three times the LLoD) in patient matrix (or equivalent). For quantitative assays, the developer can confirm the LLoQ using a surrogate panel with a concentration approximately equal to the LLoQ. In such cases, the developer should assess the concentration of the analyte stock used to build the surrogate panel using an independent reference method, if one exists. The LLoD and LLoQ should be confirmed as detailed in CLSI document EP17.²¹ CLSI document EP06²⁰ describes evaluating the linearity of quantitative assays across the AMI for genotypes and subtypes.

This testing necessitates the availability of high-concentration patient samples for each genotype and subtype. When patient samples (or isolates for rare variants) are unavailable or when obtaining high-concentration patient samples to confirm linearity of genotypes and subtypes is impossible, it may be appropriate to test plasmid, transcript, or other analyte sources to confirm the variant’s analytical performance. For microbiological tests, representative surrogate microorganisms may be selected based on genetic similarities, identifiable through *in silico* analysis tools.

7.7 Surrogate Samples for Molecular Assays

In addition to developing a surrogate sample plan (see Subchapter 5.2) and following the recommendations for surrogate sample construction described throughout this guideline, developers may need to consider additional factors when designing surrogate samples for molecular assays, including:

- Panel-based testing (eg, blending of multiple targets at different test concentrations in a single surrogate sample)
- Upstream processing steps (eg, extraction or enrichment of target sequences, preparation techniques to facilitate measurement of target sequences)

Panel-based testing enables multiple targets at different test concentrations to be assessed in a single surrogate sample. When preparing panels, the developer should carefully adjust the combinations of target concentrations so that they reflect the biological diversity of the patient population. For example, tumor mutational burden, a measurement of mutations carried by tumor cells, is a predictive biomarker being studied for potential use in determining patient response to therapy. If tumor mutational burden in the surrogate samples does not represent that of the patient sample, test performance could be confounded, because the surrogate sample could influence the result in a way that the patient sample would not. Thus, the developer needs to not only demonstrate that the surrogate analyte and matrix appropriately reflect the patient sample but show that the target concentration combinations do not interfere with or otherwise bias the assay result.

For tests with multiple targets, studies should be designed to evaluate competitive inhibition as correlated with concentration. For example, the developer should confirm that detection of one target near its LLoD is not affected by a high-normal concentration of another target (eg, for coinfection). The study should also evaluate combinations of targets that are not normally found in patient samples (eg, mutations thought to be mutually exclusive, such as epidermal growth factor receptor and KRAS mutations in NSCLC).

Upstream processing steps can also affect surrogate sample construction and composition. The treatment and processing of patient samples used for molecular testing may influence the integrity and concentration of the target sequences. For example, bisulfite conversion of DNA is the standard method for detecting unmethylated vs methylated cytosines. This conversion method involves a harsh chemical reaction that can cause significant loss of starting material and may interact unfavorably with the selected surrogate matrix. Similarly, shearing of cell-line DNA to represent cell-free DNA or circulating tumor DNA in a liquid biopsy sample can cause significant loss of amplifiable target sequences and may yield a higher-than-expected total DNA concentration in the surrogate sample.

Nucleic acids (ie, RNA and DNA) can be damaged in formalin-fixed, paraffin-embedded (FFPE) tissue samples through formalin-based cross-linkage. Depending on the fixation conditions, the amount of amplifiable nucleic acid with the target sequences may vary for a fixed input. Therefore, the developer should consider the amount of amplifiable nucleic acid available for a given input when constructing surrogate samples generated from extracted nucleic acid from more than one sample.

Finally, depending on prevalence and availability of patient samples, when constructing surrogate samples, the developer may need to prioritize either the analyte or the matrix. For some molecular tests, it may be preferable to use biological analyte in an altered matrix, rather than artificial analyte (eg, plasmid DNA) in a biological matrix. For other assays, the reverse may be true. For example, for chlamydia detection, it may be more important to emphasize the sample matrix for analytical sensitivity or artificial matrix studies, because the difficulty of sample sourcing and the large volume required for carryover or cross-contamination studies may limit the number of samples available to complete other studies. Another example is cytomegalovirus (CMV) viral load

quantitation. For low-concentration panels, the analyte source may be a patient sample, whereas tissue-derived CMV or plasmid DNA with CMV sequences may be needed for mid- and high-concentration panels. Similarly, for a liquid biopsy test, a patient sample containing a rare target may be combined with a surrogate matrix to ensure that sufficient test volume is available. For amplification-based tests that are combined with complex analysis pipelines (eg, liquid biopsy), spiking the analyte at different phases of processing can lead to dramatically different results, some of which may not mimic a patient sample.

Chapter 8

Application to Performance Studies

This chapter includes:

- Performance study principles and uses
- Factors to consider by performance study type
- Hierarchical approaches to sample selection for specific performance studies

8 Application to Performance Studies

This chapter discusses applications of surrogate samples to performance studies and factors to consider in conjunction with specific performance study objectives. Application of the surrogate sample hierarchy varies by performance study type. This chapter explains how to apply the hierarchy to specific performance studies and how to choose an appropriate surrogate sample type for the respective performance study. Each subchapter contains a table demonstrating application of the surrogate sample hierarchy to the performance study type. Table 5 provides the legend colors for these tables.

Table 5. Legend Colors for Performance Study Hierarchies^{a,4} (Modified from the Medical Device Innovation Consortium (MDIC) Clinical Diagnostics Surrogate Sample Use Working Group. *Surrogate Sample Framework: A Report of the Surrogate Sample Working Group of the Medical Device Innovation Consortium (MDIC)*. Medical Device Innovation Consortium; 2017, with permission.)

Patient sample (unspiked, individual)
This guideline and study-specific principles are used to select the surrogate sample type in a downward linear fashion. A well-thought-out rationale is needed before selection of the next surrogate sample type.
Unique properties of the assay type, study design, patient sample type, and the importance of whether altering analyte or matrix more closely mimics a patient sample in conjunction with the principles outlined in this guideline are taken into consideration. A well-thought-out rationale is needed before the next surrogate sample type is selected.
Considered only after exhausting other options.
* Remaining surrogate sample types may not be suitable for this study type.

^a Performance study hierarchies are listed from top to bottom, with dark green being most desirable and light pink least desirable.

8.1 Linearity

8.1.1 Objective

The objective of a linearity study is to determine the interval across which the average value of test replicates is proportional to the analyte concentration. The analyte concentrations should span 20% to 30% beyond the anticipated AMI (refer to CLSI document EP06²⁰). For more definitive examinations of linearity, the developer can increase either the number of replicates at each analyte concentration or the number of analyte concentrations across the target interval.

8.1.2 Principles for Surrogate Sample Use

Surrogate samples are commonly used in linearity studies to produce samples, in a representative matrix, that span the desired AMI and have a known proportional relationship. When surrogate samples are used, it is important to balance analyte and matrix integrity while minimizing matrix variability. First, the developer should determine whether the analyte for the highest sample pool is available in a patient sample. If it is not, spiking the analyte into a patient sample or patient sample pool may be necessary. Although pooled patient samples provide sufficient volume for preparing intermediate surrogate panel members and replicates for testing, the developer should consider the possibility of using a large volume of patient sample from one patient.

When the volume of a single patient sample is insufficient, to preserve analyte and matrix integrity, the developer may need to perform a study with a patient sample for a portion of the interval rather than diluting the sample. The use of individual (not pooled) patient samples to create intermediate surrogate panel members may introduce additional sample-to-sample variability. Moreover, excessively diluting patient samples to preserve analyte integrity may result in a less representative matrix. For diluted samples, the accuracy of the analytical result depends less on the matrix and more on the accuracy of the dilution and on the test method's precision at the resulting concentration. Therefore, pooled patient samples are often preferred for linearity

studies. Pooling samples minimizes sample-to-sample variability in linearity surrogate panel members and helps produce sufficient volume for study replicates. The developer can minimize any matrix effect in the intermediate surrogate panels by beginning with one surrogate pool with a relatively high analyte concentration.

For the lowest surrogate pool, obtaining a patient sample with an absence of analyte may be impossible. In this case, the developer can use an artificial matrix to create samples with analyte concentrations below the LLoD. For sex-specific analytes (eg, testosterone, prostate-specific antigen), using patient samples from the opposite sex for the low-analyte pool is acceptable. The LLoD and LLoQ may be used to establish the lowest appropriate analyte concentration (refer to CLSI document EP17²¹). For sex-specific hormone assays (eg, β hCG, luteinizing hormone), patient samples from the opposite sex can be used to create surrogate samples in linearity panels. For drug assays (eg, cocaine, sirolimus), a drug of known purity can be used to spike samples or prepare a stock solution for surrogate panels.

8.1.3 When to Use

Sufficient volume is needed for linearity studies in order to prepare dilutions across the AMI and to create the necessary number of replicates. Because it can be challenging to obtain sufficient patient sample volume for the studies suggested in CLSI document EP06,²⁰ developers can use surrogate samples to mitigate the high volume requirements.

8.1.4 How to Use

The preferred surrogate sample type for linearity studies is a high-concentration individual patient sample (A1), diluted with a blank or very low-concentration patient sample to achieve the desired analyte concentration (A2) for each linearity panel member (see Table 6). Owing to limited volume of an individual patient sample or a limited availability of patient samples with very high or very low analyte concentrations, obtaining the preferred sample may be impossible. Therefore, it is common to prepare a surrogate sample with a very high analyte concentration by spiking a patient sample with analyte (B1 preferred over E1) and then diluting it with a blank patient sample or an artificial matrix (D1 or F1) to create surrogate samples with known concentrations that span the AMI (B2 preferred over E2). Typically for E1 and E2 surrogate samples, analyte is added directly to the matrix at varying concentrations spanning the AMI, thereby creating a linearity panel.

Table 6. Linearity Study Hierarchy⁴

Linearity Sample	Sample Definition
A1	Patient sample (unspiked, individual)
A2	Pooled (unspiked, pool)
B1	Supplemented (biological spiked, individual)
B2	Pooled (biological spiked, pool)
E1	Simulated analyte (artificial spiked, individual)
E2	Simulated analyte (artificial spiked, pool)
C1	Simulated matrix (unspiked, individual)
C2	Simulated matrix (unspiked, pool)
D1	Simulated matrix (biological spiked, individual)
D2	Simulated matrix (biological spiked, pool)
F1	Simulated analyte/matrix (artificial spiked, individual)
F2	Simulated analyte/matrix (artificial spiked, pool)
H	Simulated matrix (biological spiked, artificial)
I	Simulated matrix (artificial spiked, artificial)
G	Simulated matrix (unspiked, artificial)

8.2 Analytical Specificity

8.2.1 Objective

Analytical specificity (ie, interference, cross-reactivity, competitive inhibition) studies evaluate whether materials present in patient samples change the results of a test. The objective of these studies is to demonstrate test performance in the presence of potentially interfering substances (refer to CLSI document EP07²⁷). Frequently, these assessments test both analyte-positive and blank samples.

8.2.2 Principles for Surrogate Sample Use

Because analytical specificity assesses the analyte at specific concentrations in combination with potentially interfering or cross-reactive substances at specific concentrations, surrogate samples enable testing of combinations that are rare or nonexistent in patient populations (eg, extremely high levels of an interferent) or are difficult to obtain. In order to represent the patient sample matrix, it is important to maintain individual sample variability. When using surrogate samples for analytical specificity tests, the developer needs to consider the extraction method (if part of the prescribed assay methodology). Surrogate samples should be processed through all sample workflow steps, including extractions, to represent patient sample preparation and testing.

8.2.3 When to Use

Surrogate samples can be used to control the amount or concentration of potential interferences, thereby making it possible to quantitate the magnitude of any change in result as a function of the interferent.

8.2.4 How to Use

The preferred sample type for analytical specificity studies is an unspiked individual patient sample (A1), but only when patient samples are unmanipulated and when both the analyte and the potential interferent can be quantitated by highly specific methodologies (eg, a recognized reference method). If this sample type is insufficient, the developer should choose the preferred surrogate sample, following the order shown in Table 7. Although individual patient samples are preferred, pooling samples may be necessary if unknown

and/or significant levels of potentially interfering materials are a concern. Creating surrogate sample pools with concentrations at the MDLs (for quantitative assays) or two to three times the LLoD (for qualitative assays) or pools that span the AMI may necessitate pooling patient samples from the intended-use population. When using surrogate sample pools, the developer should be careful, because potential cross-reactants and interferents present in individuals with that disease state may be masked by pooling. When the analyte is an exogenous material (eg, for qualitative infectious diseases assays) and blank patient samples are needed (eg, negative swab samples), using a simulated artificial matrix (G) ensures that the samples are truly negative.

Table 7. Analytical Specificity Study Hierarchy⁴

Analytical Specificity Sample	Sample Definition
A1	Patient sample (unspiked, individual)
A2	Pooled (unspiked, pool)
B1	Supplemented (biological spiked, individual)
B2	Pooled (biological spiked, pool)
E1	Simulated analyte (artificial spiked, individual)
E2	Simulated analyte (artificial spiked, pool)
C1	Simulated matrix (unspiked, individual)
C2	Simulated matrix (unspiked, pool)
D1	Simulated matrix (biological spiked, individual)
D2	Simulated matrix (biological spiked, pool)
F1	Simulated analyte/matrix (artificial spiked, individual)
F2	Simulated analyte/matrix (artificial spiked, pool)
G	Simulated matrix (unspiked, artificial)
H	Simulated matrix (biological spiked, artificial)
I	Simulated matrix (artificial spiked, artificial)

8.2.5 Additional Considerations

Potentially interfering substances can be introduced through patient sample preparation. For example, the fixation process for FFPE tissue may introduce materials that interfere with test performance or otherwise diminish sample integrity. In some cases, the developer can evaluate the potential interference by comparing patient samples and surrogate samples processed using different procedural conditions.

The addition of potentially interfering material may not properly reflect physiological conditions. For example, adding a drug (diluted or undiluted) to a surrogate sample may not yield the biologically active metabolite (eg, omission of the active metabolite N-acetylprocainamide during development of a test for procainamide). Alternately, the binding constants for the potentially interfering material may influence test results (eg, use of caffeine or theophylline for measurement of free vs protein-bound bilirubin). For these surrogate samples, the developer may need to refer to the literature for recommended practices, based on historical or otherwise reasonable expectations, for using the potentially interfering material in surrogate samples (refer to CLSI document EP37⁴³).

For some potential interferents, an appropriate surrogate sample may not exist. For example, necrotic tissue in an FFPE tissue sample cannot be simulated with a surrogate sample. If no suitable surrogate sample exists, the developer should use patient samples that contain the potentially interfering materials. During analytical

specificity testing for infectious diseases, surrogate samples created to have concentrations at the MDLs may not be reactive for the same antigens that the test is designed to detect. This type of surrogate sample may not be appropriate.

8.3 Precision

8.3.1 Objective

Precision studies evaluate different components of variance, such as within-run (ie, repeatability), between-run, between-day, and between-lot. They may also assess other variables, such as site-to-site variability. The objective of these studies is to establish the level of agreement among repeated measurements on the same or similar sample(s) under different conditions (eg, operator, instrument, time, location, conditions of use). CLSI document EP05²⁶ provides full details on the scope and experimental design of precision studies. Appendixes C, D, and E include example precision studies.

8.3.2 Principles for Surrogate Sample Use

Surrogate samples are widely used in precision studies to obtain sufficient volume and ensure the stability of the sample that contains an appropriate analyte concentration so that the required number of replicate measurements can be taken over a specified time frame. The preferred surrogate sample type for precision studies is pooled patient samples or spiked individual patient samples. If the pooled patient samples do not contain the desired amount of analyte, the analyte concentration can be increased by spiking or decreased by dilution with a blank patient sample pool.

8.3.3 When and How to Use

Owing to the necessity for high sample volume and sample stability, surrogate samples are often needed for precision testing. Using surrogate samples enables the developer to obtain sufficient sample volume, with known analyte concentrations across the AMI. Surrogate samples can be prepared in several ways, including from individual or pooled patient samples, either unspiked or spiked with analyte. As shown in Table 8, the preferred sample type is an unspiked individual patient sample (A1). When this sample type is insufficient, an unspiked pooled patient sample (A2) is preferred, followed by an individual patient sample spiked with a different patient sample (B1), then by a pooled patient sample spiked to achieve the desired analyte concentration (B2).

If a high-concentration patient sample is unavailable, the developer needs to decide whether it is more important for the surrogate sample to have an unaltered analyte or an unaltered matrix. When an unaltered matrix is more important, it may be acceptable to spike the individual or pooled patient sample with an appropriate stock material until the desired analyte concentration is reached (E1 preferred over E2). For multiplexed tests (eg, tests with a single signature score), creating samples with multiple targets may be useful (see Subchapter 7.5). In such cases, using less preferred options, such as an altered sample matrix (C1, C2, D1, D2, F1, F2) or an artificial matrix such as the assay diluent (H, I, G), may sometimes be appropriate. When an unaltered analyte is more important, after B2, the next best options are C1, D1, C2, and finally D2. For example, for molecular tests, altered sample matrix that contains biological analyte (C1, D1) may be preferable to biological matrix spiked with artificial nucleic acid (E1).

Table 8. Precision Study Hierarchy⁴

Precision Sample	Sample Definition
A1	Patient sample (unspiked, individual)
A2	Pooled (unspiked, pool)
B1	Supplemented (biological spiked, individual)
B2	Pooled (biological spiked, pool)
E1	Simulated analyte (artificial spiked, individual)
E2	Simulated analyte (artificial spiked, pool)
C1	Simulated matrix (unspiked, individual)
C2	Simulated matrix (unspiked, pool)
D1	Simulated matrix (biological spiked, individual)
D2	Simulated matrix (biological spiked, pool)
F1	Simulated analyte/matrix (artificial spiked, individual)
F2	Simulated analyte/matrix (artificial spiked, pool)
H	Simulated matrix (biological spiked, artificial)
I	Simulated matrix (artificial spiked, artificial)
G	Simulated matrix (unspiked, artificial)

8.3.4 Additional Considerations

When selecting surrogate sample types, the developer should consider using individual spiked patient samples to better represent variability among samples. For example, spiking analyte into individual patient samples retains the matrix variability. Although comprehensive precision studies require large numbers of patient samples, replacing patient samples with surrogate samples may lead to an incorrect conclusion on test performance. When possible, the developer may need to supplement small-scale precision studies (that use patient samples) with a larger-scale study (that uses surrogate samples). CLSI recognizes that pooled patient samples are typically used for precision studies and that spiking or diluting samples is necessary to achieve appropriate analyte concentrations.

8.4 Detection Capability

8.4.1 Objective

Detection capability studies assess the ability of a test to detect and accurately determine the analyte's presence at low concentrations (based on the test's performance limits), when low analyte concentrations are clinically significant (refer to CLSI document EP17²¹). These studies are used, depending on the type of test, to establish the LoB (see Table 9), LLoD (see Table 10), and/or LLoQ (see Table 11). These assessments, usually the first validation studies conducted, are used to determine the appropriate analyte concentration for other studies (eg, interfering substances, precision).

8.4.2 Principles for Surrogate Sample Use

To assess detection capability, it is best to use multiple independent blank and/or low-concentration patient samples, which accounts for matrix variability among samples. In some cases (eg, if the developer needs a large volume of samples or samples with desired analyte concentrations), using surrogate samples is acceptable, provided they perform similarly to patient samples in the test method. The matrix selected should closely represent individual patient sample matrix. Moreover, the selected sample should not contain extraneous

substances that could influence the test results. Such exogenous substances may introduce bias, which could affect the study's ability to establish the "true" detection limit of a test or test system. If the matrix is artificially prepared, the surrogate sample's characteristics should be clearly defined, and the surrogate matrix should resemble that of an appropriate patient sample.

8.4.3 When to Use

Surrogate samples can be effectively used to determine analytical detection limits when sufficient volumes of patient samples with known concentrations (including very low or blank) of the analyte(s) of interest are difficult to find in a clinical population. Using surrogate samples may also be appropriate when international standards of the analyte(s) are available (see Additional Resources).

8.4.4 How to Use

To establish or verify the LLoD and LLoQ for a study, the developer should prepare surrogate samples with the desired analyte concentrations by diluting positive patient samples or spiking biological or artificial analyte into blank patient samples. Preferably, an analyte(s) that is spiked into the sample matrix should have a high (natural or artificial) concentration, in order to limit any dilution effects or potential errors caused by spiking too much material. For example, when evaluating the LLoD or LLoQ for β hCG, the developer can spike β hCG-free male sample pools with a low-concentration β hCG female sample pool.

The preferred sample type is an unspiked individual patient sample (A1) known to be blank or below the intended detection limit for the intended analyte(s). If this sample type is insufficient, the developer should choose the preferred surrogate sample, following the orders shown in Tables 9, 10, and 11. This order of preference is based on the principle of maintaining individual sample variability to better represent patient sample performance. If the needed volume cannot be obtained from an individual patient sample, the next best option is to pool as few patient samples as possible, thus preserving the matrix's biological composition. The test designs recommended in CLSI document EP17²¹ can help determine whether it is better to use multiple pools composed of small numbers of patient samples or a single pool composed of a large number of patient samples. Additionally, CLSI document EP17²¹ can help the developer estimate the sample volume needed for each anticipated study and decide which statistical analysis to use.

If it is difficult to obtain a sufficient volume of known blank patient sample, the developer can use an artificial matrix that closely resembles the patient sample matrix. When identifying individual patient samples that are negative for the target analyte is impossible, a blank sample can be used to determine the LoB. It does not need to contain the target analyte (ie, it can be a buffer). Alternatively, the sample may have a concentration at least an order of magnitude lower than the lowest concentration of interest.

When preparing individual patient samples that are negative for the target analyte is possible, the developer should add analyte to the matrix directly from a high-concentration stock or from a patient sample that contains the intended analyte(s) (if quantitatively measured). This surrogate sample can be used to make serial dilutions or admixtures, thereby creating samples with the concentrations needed to determine the LLoD. Preferably, dilutions should be made with the unspiked blank patient sample matrix.

International reference standards may be used as a source of analyte for surrogate samples (see CLSI document EP05²⁶). In some cases, standard reference material can be used even if it has a different matrix (eg, National Institute of Standards and Technology reference material in a serum matrix for use in a urine assay), provided that the matrix does not affect the assay's bias or precision (see Additional Resources). The developer can add analyte to individual patient samples that are negative for the target analyte, using the high-concentration stock prepared from the international standard or using serial dilutions made from the high-concentration stock, to create samples with the concentrations needed to determine the LLoD. The amount of international

standard (in a different matrix) should be minimized to prevent significant changes to the surrogate sample matrix. Preferably, dilutions should be made with the unspiked blank patient sample matrix. The advantage of using the international reference standard is that the starting concentration of the source stock is known. The developer should assess the source material from which the international standard was prepared, because some international standard reference materials are prepared using a recombinant source that may not perform equivalently to patient samples.

When blank patient samples are used, it is important to reduce the bias associated with their variability. Samples obtained from different patients in a clinical setting may contain different levels of endogenous and/or exogenous components based on the patient's genetic and/or disease state, as well as different substances (eg, medications) used to manage the disease. Although the presence of the endogenous elements can help establish analytical sensitivity in the intended-use population, in excess, these substances can create a matrix effect, which may generate bias in the study results. Patient samples should be free of obvious interferents that may generate bias. Nonetheless, surrogate samples should adequately reflect the complexities in the desired sample type, and the matrix should not be artificially simplified.

Table 9. LoB Study Hierarchy⁴

LoB Sample	Sample Definition
A1	Patient sample (unspiked, individual)
A2	Pooled (unspiked, pool)
C1	Simulated matrix (unspiked, individual)
C2	Simulated matrix (unspiked, pool)
G	Simulated matrix (unspiked, artificial)
*	Remaining surrogate sample types may not be suitable for this study type.

Abbreviation: LoB, limit of blank.

Table 10. LLoD Study Hierarchy⁴

LLoD Sample	Sample Definition
A1	Patient sample (unspiked, individual)
B1	Supplemented (biological spiked, individual)
E1	Simulated analyte (artificial spiked, individual)
A2	Pooled (unspiked, pool)
B2	Pooled (biological spiked, pool)
E2	Simulated analyte (artificial spiked, pool)
C1	Simulated matrix (unspiked, individual)
D1	Simulated matrix (biological spiked, individual)
F1	Simulated analyte/matrix (artificial spiked, individual)
C2	Simulated matrix (unspiked, pool)
D2	Simulated matrix (biological spiked, pool)
F2	Simulated analyte/matrix (artificial spiked, pool)
H	Simulated matrix (biological spiked, artificial)
I	Simulated matrix (artificial spiked, artificial)
*	Remaining surrogate sample types may not be suitable for this study type.

Abbreviation: LLoD, lower limit of detection.

Table 11. LLoQ Study Hierarchy⁴

LLoQ Sample	Sample Definition
A1	Patient sample (unspiked, individual)
A2	Pooled (unspiked, pool)
B2	Pooled (biological spiked, pool)
E2	Simulated analyte (artificial spiked, pool)
D2	Simulated matrix (biological spiked, pool)
F2	Simulated analyte/matrix (artificial spiked, pool)
H	Simulated matrix (biological spiked, artificial)
I	Simulated matrix (artificial spiked, artificial)
*	Remaining surrogate sample types may not be suitable for this study type.

Abbreviation: LLoQ, lower limit of quantitation.

8.4.5 Additional Considerations

Preparing surrogate samples for use in LLoD studies may be challenging, because the preparation method may introduce bias. For example, reliably spiking very low amounts of analyte into samples may exceed the capability of the tools. Therefore, the developer may need to create serial dilutions, whereby a high analyte concentration is added to a matrix, which is then diluted multiple times to achieve the appropriate concentration. This technique is used for materials in solution, not materials in suspension, because homogeneity cannot be ensured for suspensions. The developer needs to consider the error associated with multiple dilutions when setting specifications for accuracy and bias assessment.

Some sample types may create additional challenges when special methodologies are needed to assess the detection limits. For example, when FFPE tissue is the sample type used for cancer diagnostic tests, insertions or deletions of one or more nucleotides may be needed for surrogate samples. For swab specimens, it may be impossible to add the analyte directly to the swab. In such cases, the developer may need to add the analyte to an appropriate liquid matrix. Various dilutions are then coated onto the swab.

The developer needs to consider the stability of the surrogate panels manufactured for LoB, LLoD, or LLoQ studies (see Subchapter 6.4). Excessive stripping of the patient sample matrix may result in concentrations lower than the lowest value to which the test is calibrated (ie, zero calibrator). Thus, the developer would need to justify removal of multiple test concentrations that yielded negative values against the background (ie, zero calibrator). Spiking this matrix to be equal to or higher than the calibrator zero may be necessary. In some situations, the developer needs to be able to create reproducible blank sample matrix pools, so that the blank matrix can be used for other validation studies.

8.5 Matrix Comparison

8.5.1 Objective

The objective of a matrix comparison study is to assess the comparability of the primary patient sample type and a candidate patient sample type when the same test method is used. This study involves establishing the magnitude of bias in analyte measurement or detection in two similar matrixes (eg, serum and plasma, or the same specimen type collected with different anticoagulants or collection devices) or the clinical suitability of using two dissimilar matrixes (eg, serum and urine). For similar matrixes, in which the test method's performance characteristics and MDLs are similar, the developer can assess suitability by verifying the candidate sample type's performance against the established performance claims for the primary sample type. For dissimilar

matrixes, in which the test method's performance characteristics and MDLs are different, the developer may need to establish the analytical and clinical characteristics of each sample type. Matrix comparison studies can be performed for quantitative, semiquantitative, and qualitative test methods. For a quantitative or semiquantitative test, the developer needs to determine whether there is a systematic difference between the two matrix types across the AMI and at the MDLs. For a qualitative test, the developer should estimate the PPA and NPA of the test results between the two matrix types.

8.5.2 Principles for Surrogate Sample Use

Surrogate samples may be used to facilitate matrix comparison studies by providing sufficient sample volume to conduct tests or to obtain targeted analyte concentrations. The stability of surrogate matrixes should be considered.

8.5.3 When to Use

8.5.3.1 Coverage of the Analytical Measuring Interval

For a quantitative test, procuring patient samples with high or low concentrations may be difficult. In this case, surrogate samples can be used to cover the AMI.

8.5.3.2 Samples With Concentrations Close to the Medical Decision Levels or Cutoff

Samples with concentrations close to the MDLs or cutoff are needed for matrix comparison studies. If procuring such patient samples is difficult, surrogate samples can be used.

8.5.3.3 Insufficient Samples With Positive Test Results

For a qualitative test, estimating the PPA requires a certain number of patient samples with positive test results. If it is difficult or unethical to procure patient samples with positive test results (eg, samples positive for drugs of abuse), surrogate samples can be included.

8.5.3.4 Invasive Sampling Methods

Some specimens (eg, CSF) can only be collected through invasive procedures that pose a high risk of harm to the patient. Depending on the benefits vs risks of obtaining blank patient samples, surrogate samples may be appropriate for a defined population.

8.5.4 How to Use

When an unspiked individual patient sample (A1) is not available, the developer should choose the preferred surrogate sample, following the order shown in Table 12, which prioritizes surrogate sample types that preserve the sample matrix. Surrogate samples composed of individual patient samples supplemented with biological analyte (B1) are preferred over samples supplemented with artificial analyte (E1). When individual surrogate samples (B1 and E1) cannot be used, the developer may consider using pooled surrogate samples (A2, B2, E2). When pooling is necessary, the contributions from each individual patient sample for the candidate sample type should be proportional to those for the primary sample type. For example, if serum from patient A constitutes 60% of a serum pool, plasma from patient A should also compose 60% of the plasma pool. Other conditions may apply when a reference method validated for the primary matrix is used to compare the acceptability of alternative matrixes (see CLSI document EP35²⁴). Generally, when patient samples are supplemented with surrogate samples, the panel of surrogate samples should include concentrations that overlap those of several patient samples, thereby enabling the performance of patient and surrogate samples to be directly compared. Matrix comparison studies should be conducted within the time frame for which stability has been established for both sample types.

For quantitative tests, when supplementing surrogate samples of the candidate sample type with analyte from patient samples of the primary sample type, the developer should add the lowest practical amount of analyte in order to minimize the effect on the candidate sample matrix. For example, a patient serum sample (ie, the primary sample type) is needed to create a surrogate plasma sample (ie, the candidate sample type). The patient sample needs to have a high analyte concentration to minimize the amount of serum introduced into the plasma matrix. After ensuring that the serum will form a very small percentage of the plasma sample, the developer should spike a low volume of serum to create a surrogate plasma sample.

For qualitative tests, when using surrogate samples, the developer should preserve samples with concentrations close to the cutoff for the matrix comparison study. The C5 to C95 interval (or LLoD) of the qualitative test should be assessed for both matrixes. Using two surrogate sample panels with similar analyte concentrations can enable direct comparison of qualitative test responses or statistical parameters (eg, PPA and NPA). Moreover, increasing replicate testing around the defined cutoff (eg, LLoD, clinical sensitivity and clinical specificity) can increase confidence that both matrixes are associated with a similar response.

Table 12. Matrix Comparison Study Hierarchy⁴

Matrix Comparison Sample	Sample Definition
A1	Patient sample (unspiked, individual)
B1	Supplemented (biological spiked, individual)
E1	Simulated analyte (artificial spiked, individual)
A2	Pooled (unspiked, pool)
B2	Pooled (biological spiked, pool)
E2	Simulated analyte (artificial spiked, pool)
*	Remaining surrogate sample types may not be suitable for this study type.

8.6 Method Comparison

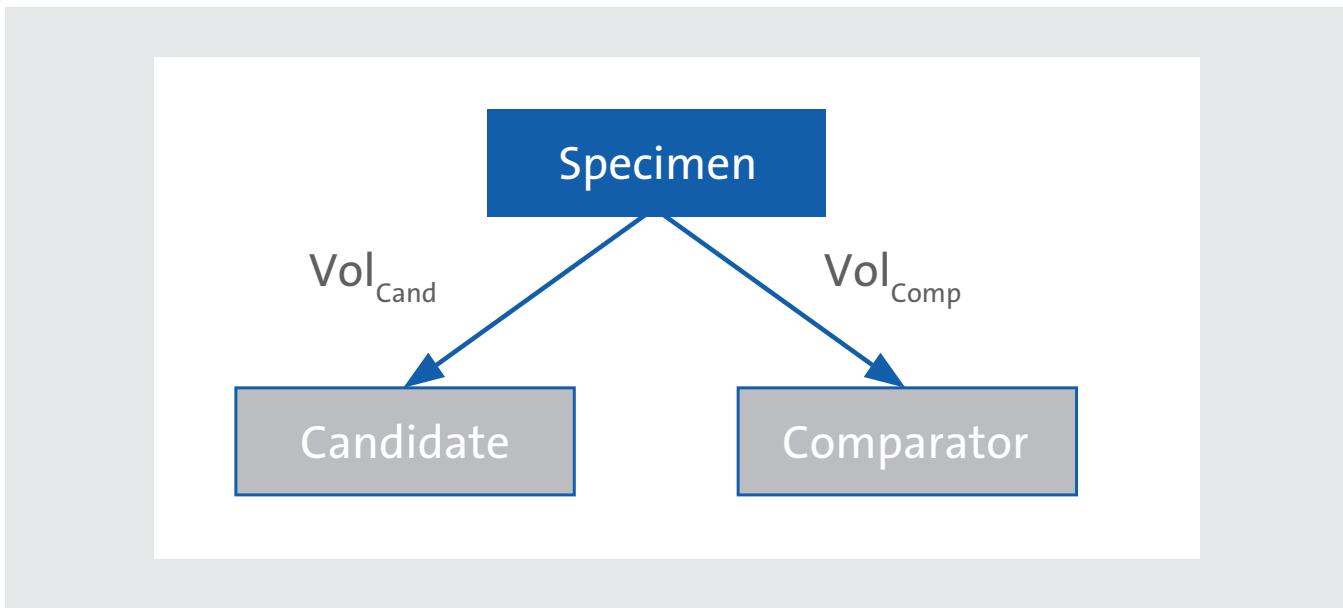
8.6.1 Objective

Method comparison studies assess the comparability of a candidate method and an established method, called the “comparative method.” The objective of these studies is to determine whether the candidate method is analytically and clinically equivalent to the comparative method. For an example of a method comparison study, see Appendix F.

This subchapter describes the use of surrogate samples for comparative and candidate methods that involve the same matrix. If the matrixes differ by method (eg, the candidate method uses whole blood and the comparator uses plasma), surrogate sample preparation and results interpretation can become much more complex, and the developer may need to provide an additional rationale and justification for use. Study designs that use multiple matrixes should be an exception, not routine practice.

8.6.2 Principles for Surrogate Sample Use

Generally, each sample should have sufficient volume for at least one measurement for each method. This design, described as “split sample” (refer to CLSI document EP12¹⁸), is shown in Figure 6.



Symbols: Vol_{Cand} , sample volume measured with candidate method; Vol_{Comp} , sample volume measured with comparative method.

Figure 6. Comparison Study Split-Sample Design

To the extent possible, samples should reflect biological variability. When pooling patient samples for method comparison studies, the developer should:

- To best represent the variability of patient samples, create “minipools” by mixing as few samples as possible.
- To maintain the biological composition of the minipools:
 - If there is sufficient volume, use no more than two samples to create the minipool.
 - If there are enough samples, use each sample for only one minipool.
- To minimize the increase in variability, pool samples with similar characteristics (eg, matrix, disease state).

An exception to these general principles is a patient sample with an unusual characteristic (eg, a superhyperlipidemic sample), which may be used to spike several pooled patient samples. In this case, the developer should determine whether and to what degree spiking the exceptional sample will decrease the biological variability of the samples. Figure 7 illustrates the application of the first two principles in design B. Design A, which does not follow these principles, results in less diverse minipools and is not recommended.

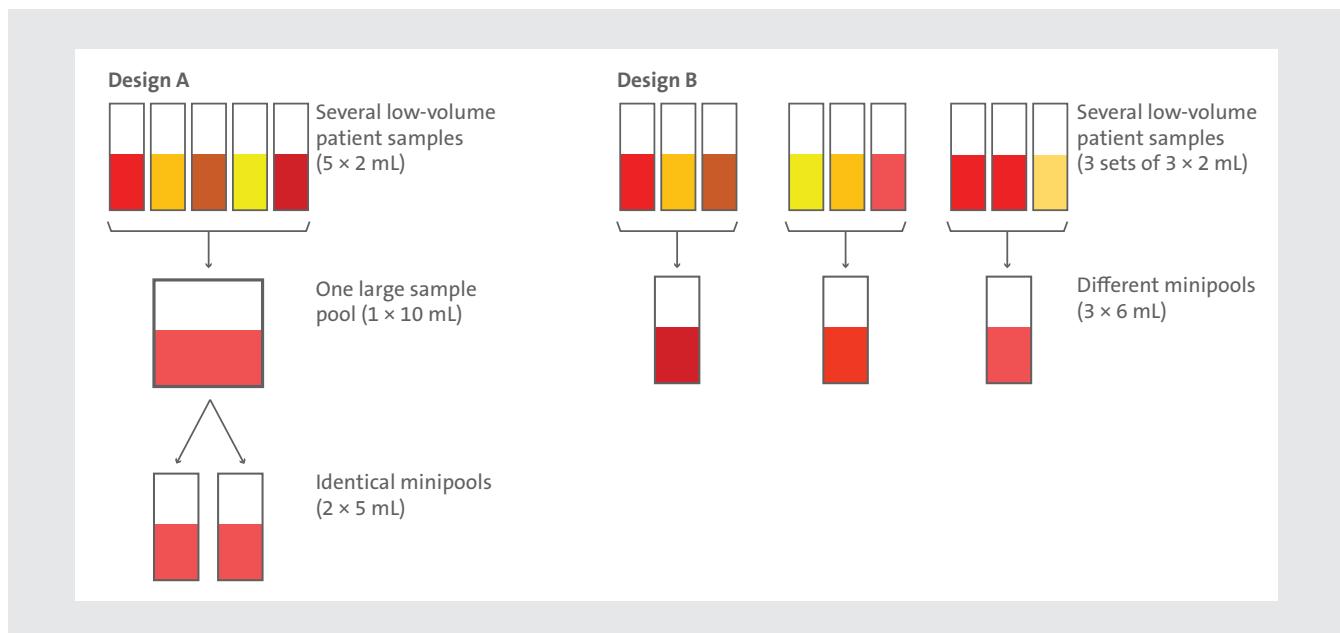


Figure 7. Different Pooling Designs. The needed volume for testing both the comparative and candidate methods is 5 mL. Each patient sample has a volume of 2 mL.

When supplementing samples for method comparison studies, the developer should:

- Use stocks with the highest available concentrations to spike samples, in order to minimize the amount of spiking material introduced into the sample.
- Use patient samples with naturally low analyte concentrations or blank patient samples, in order to minimize the amount of diluent.

8.6.3 When to Use

8.6.3.1 Insufficient Sample Volume

When the volume of each individual patient sample is insufficient to test both the comparative and candidate methods, minipools of patient samples can be used.

8.6.3.2 Coverage of the Analytical Measuring Interval

For a quantitative test, procuring patient samples with high or low concentrations may be difficult. In this case, surrogate samples can be used to cover the AMI.

8.6.3.3 Samples With Concentrations Close to the Medical Decision Levels

For quantitative tests, a sufficient number of patient samples with concentrations close to the MDLs are needed to meaningfully estimate systematic differences or bias at the MDLs. When procuring such samples is difficult, surrogate samples can be used.

8.6.3.4 Insufficient Samples With Positive Test Results

For a qualitative test, estimating the PPA requires a certain number of patient samples with positive test results. If it is difficult to procure patient samples with positive test results, surrogate samples can be included. Only certain types of studies may include surrogate samples. Although surrogate samples can be useful for relative comparisons of qualitative methods (ie, PPA and NPA), they cannot be used to establish the candidate method's clinical sensitivity or clinical specificity.

8.6.4 How to Use

When unspiked individual patient samples (A1) are unavailable, the preferred surrogate sample type is individual patient samples, spiked with either a different patient sample or with a high-concentration pooled patient sample, to achieve the desired analyte concentration (B1). The developer should choose the preferred surrogate sample, following the order shown in Table 13. If individual patient samples are not available, the next best option is unspiked pooled patient samples, followed by pooled patient samples spiked with either a different patient sample or with a high-concentration pooled patient sample, to achieve the desired analyte concentration (A2 preferred over B2). Pools of samples with similar characteristics (eg, concentration, disease state) are created. If a high-concentration patient sample is unavailable, the developer may use an appropriate stock material to spike the individual or pooled patient sample to the desired analyte concentration (E1 preferred over E2). Depending on the sample type used for the test, using less preferred options, such as an altered matrix (C1, C2, D1, D2, F1, F2) or an artificial matrix such as the assay diluent (H, I), may sometimes be appropriate.

Table 13. Method Comparison Study Hierarchy⁴

Method Comparison Sample	Sample Definition
A1	Patient sample (unspiked, individual)
B1	Supplemented (biological spiked, individual)
A2	Pooled (unspiked, pool)
B2	Pooled (biological spiked, pool)
E1	Simulated analyte (artificial spiked, individual)
E2	Simulated analyte (artificial spiked, pool)
C1	Simulated matrix (unspiked, individual)
C2	Simulated matrix (unspiked, pool)
D1	Simulated matrix (biological spiked, individual)
D2	Simulated matrix (biological spiked, pool)
F1	Simulated analyte/matrix (artificial spiked, individual)
F2	Simulated analyte/matrix (artificial spiked, pool)
H	Simulated matrix (biological spiked, artificial)
I	Simulated matrix (artificial spiked, artificial)
*	Remaining surrogate sample types may not be suitable for this study type.

8.6.5 Additional Considerations

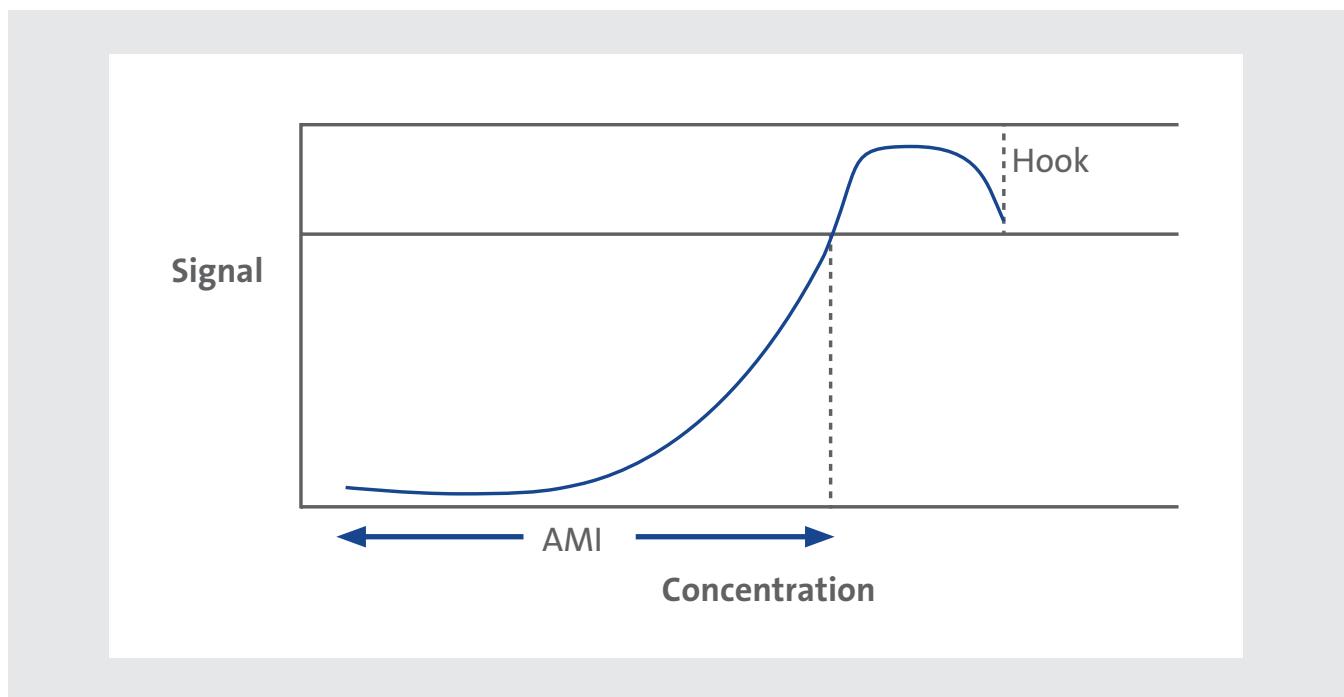
For qualitative tests (ie, tests with either positive or negative outputs), the developer can demonstrate the candidate method's analytical performance by the PPA and NPA, which are calculated by comparing results of the comparative vs candidate tests. When using surrogate samples for the method comparison study, the developer should preserve samples with concentrations close to the cutoff (if known), in order to obtain unbiased estimates of the PPA and NPA (refer to CLSI document EP09²²). The developer can perform separate statistical analyses for patient samples and for surrogate samples to determine whether the data for the overall analysis can be combined. The PPAs and NPAs should be similar for patient and surrogate samples. If this is the case, the developer can perform an analysis for all patient and surrogate samples combined, as applicable.

For quantitative tests, one performance parameter in the method comparison study is estimates of bias (ie, systematic differences) at the MDLs. Estimates of bias at the MDLs should be similar for patient and surrogate samples. If this is the case, the developer can perform an appropriate regression analysis for all patient and surrogate samples combined (see Subchapter 7.2.1).

8.7 Hook Effect

8.7.1 Objective

The objective of a hook effect study is to determine whether a high-dose hook effect (ie, prozone effect) in an immunometric (ie, sandwich) assay exists and whether the effect is clinically significant. The high-dose hook effect refers to the measured analyte concentration displaying a significantly lower signal (eg, absorbance or fluorescence) than predicted for that concentration. A “hook” is observed within the calibration range when signal is plotted vs analyte concentration (see Figure 8).



Abbreviation: AMI, analytical measuring interval.

Figure 8. Hook Effect. A false low result is generated for a sample whose concentration is above the AMI.

8.7.2 Principles for Surrogate Sample Use

Surrogate samples are commonly used in high-dose hook effect studies to produce a sample with a concentration above the established AMI (see CLSI document EP34²⁸). When using surrogate samples, the developer should balance matrix and analyte integrity while minimizing matrix variability, because the objective is to assess the analytical response to analyte concentration. Because only a few diluted samples and replicates are needed, a single very high-concentration patient sample is sufficient, if available. Otherwise, a pooled patient sample may be adequate. Although a pooled patient sample provides sufficient volume to prepare intermediate-concentration surrogate samples (ie, a panel) for replicate testing, an individual patient sample is preferred. Excessively diluting the sample to preserve analyte integrity may make the matrix less representative of individual patient sample matrix. Because it may be difficult to find a patient sample whose concentration is equal to or near the highest clinically anticipated concentration, surrogate samples are commonly used for hook effect studies.

8.7.3 When to Use

Because patient samples with concentrations above the AMI can be difficult to obtain, surrogate samples are often used.

8.7.4 How to Use

As shown in Table 14, the preferred surrogate sample type is an unspiked individual patient sample (A1) whose concentration is above the established AMI and close to the maximum clinically anticipated concentration, diluted with a blank or very low-concentration patient sample to achieve the desired analyte concentration for each panel member. Owing to a limited volume of individual patient samples or the rarity of a patient sample whose concentration is above the established AMI, obtaining this sample type may be impossible. Therefore, it is common to prepare a very high-concentration sample by spiking a sample with analyte (B1 preferred over E1) and then diluting it with a blank patient sample or a patient sample whose concentration is well within the AMI. The concentration of the highest spiked sample can be verified by several methods (eg, independent instrument system, dilutions of the initial spiked sample, gravimetric addition of a value-assigned concentrate) (refer to CLSI document EP34²⁸).

The objective of preparing this panel is to provide several samples with varying concentrations, from very high to within the AMI. Thus, the developer can establish a reasonable estimate of the analyte concentration for samples that are expected to produce false low results. For example, to create a high-concentration surrogate sample to study the hook effect for an ovarian tumor marker assay, the developer could use an antigen composed of concentrated cell culture supernatant from a human ovarian carcinoma cell line and spike it into a native sample. As another example, to demonstrate that a procalcitonin assay is free of hook effects up to a specific concentration, the developer could create one high-concentration sample (eg, serum spiked with recombinant procalcitonin stock) and generate multiple procalcitonin concentrations throughout the AMI and at MDLs by diluting the samples with the test procedure diluent.

Table 14. Hook Effect Study Hierarchy⁴

Hook Effect Sample	Sample Definition
A1	Patient sample (unspiked, individual)
B1	Supplemented (biological spiked, individual)
E1	Simulated analyte (artificial spiked, individual)
B2	Pooled (biological spiked, pool)
E2	Simulated analyte (artificial spiked, pool)
*	Remaining surrogate sample types may not be suitable for this study type.

8.8 Sample Stability

8.8.1 Objective

A stable sample maintains its measured characteristic(s) within acceptable limits over an established time period. The objective of a sample stability study is to assess the magnitude of analyte degradation, through the sample's measured characteristic(s), over time when the sample is stored or handled under defined conditions (refer to CLSI document EP35²⁴; CLSI document EP25³² may also be useful).

8.8.2 Principles for Surrogate Sample Use

A developer can identify bias or variability changes that arise from sample liability by assessing analyte stability in a known matrix (eg, patient or surrogate matrix) or in a sample that has been manipulated by additional processing (eg, extraction). The use of surrogate samples to establish sample stability is not well documented in the literature. When patient samples are not available, surrogate samples composed of a well-characterized matrix of known stability may facilitate assessment of analyte stability when tested with reagents of known stability.

8.8.3 When to Use

Generally, surrogate samples may be used when:

- The patient samples are of insufficient volume for the duration of sample stability testing.
- The availability of patient samples with a known analyte concentration in a specific interval(s) is limited.
- Emerging threat analytes or analytes associated with rare disease states are present.

To avoid introducing bias, the developer should consider the stability of the selected matrix. The same considerations and recommended hierarchy apply for samples that will be stored for future use after undergoing procedural manipulation (eg, extraction).

8.8.4 How to Use

As shown in Table 15, because analyte degradation can differ by matrix, the preferred sample type for sample stability studies is an unspiked individual patient sample (A1) of sufficient volume containing a known analyte concentration within a specified interval (eg, at the MDL, within the AMI [for a quantitative assay], at two to three times the LLoD [for a qualitative assay]). In some cases, patient samples characterized to this extent are difficult to find. Therefore, surrogate samples are used instead. To closely represent the matrix characteristics of a patient sample, the preferred surrogate sample is a blank individual patient sample of sufficient volume spiked with biological analyte (B1). Regardless of the sample type used, the developer should consider the potential for analyte degradation. The in-matrix stability of an artificial analyte may differ from that of the corresponding biological analyte. When the patient samples are of insufficient volume for the duration of sample stability testing, the next best option to pool as few patient samples as possible (A2 preferred over B2). With proper justification, artificial analytes may be acceptable for use (E1 preferred over E2). However, using altered or artificial matrixes for sample stability studies is not recommended.

The developer should consider how well the surrogate matrix resembles patient sample matrix in terms of viscosity and other key elements (eg, natural components of the sample). Before examination, surrogate samples should be processed with the same preexamination steps as patient specimens, so that spiked analyte is processed equivalently to the analyte in patient specimens. Generally, each level of the hierarchy guides the developer to the best options for selecting surrogate samples, based on the scientific rationale. This is particularly true when surrogate samples are used to resolve the scientific challenges of certain patient samples (eg, the low volume of whole blood or CSF samples).

Table 15. Sample Stability Study Hierarchy⁴

Sample Stability Sample	Sample Definition
A1	Patient sample (unspiked, individual)
B1	Supplemented (biological spiked, individual)
A2	Pooled (unspiked, pool)
B2	Pooled (biological spiked, pool)
E1	Simulated analyte (artificial spiked, individual)
E2	Simulated analyte (artificial spiked, pool)
*	Remaining surrogate sample types may not be suitable for this study type.

8.9 Carryover

8.9.1 Objective

The objective of a carryover study is to demonstrate that carryover and cross-contamination will not occur when automated liquid handling systems are used with patient samples.^{44,45} These phenomena may contribute to a false-negative or false-positive result, incorrect concentration, indeterminate result, or no result. These studies can be smaller if the system design prevents the possibility of carryover or cross-contamination. For an example of a carryover study, see Appendix G.

8.9.2 Principles for Surrogate Sample Use

The surrogate matrix should closely represent patient matrix. If the matrix used was derived from nonhuman samples (eg, horse blood), the surrogate sample's properties may not properly represent the characteristics of the intended patient sample type that may contribute to carryover. For example, owing to its viscosity, the surrogate sample may not correctly reflect the potential for patient sample and/or analyte carryover effects.

8.9.3 When to Use

When there are not enough patient samples with known, high analyte concentrations in the clinical population to assess the potential for carryover and cross-contamination, surrogate samples can be used. Individual or pooled patient samples spiked with a high analyte concentration should be used to simulate high-positive (ie, high-concentration) samples (eg, at a concentration at least equal to the highest clinically anticipated concentration). Individual patient samples known to be blank are also used. Developers can use blank surrogate samples when the volume of individual blank patient samples is insufficient and they cannot ensure that pooled patient samples are analyte free.

8.9.4 How to Use

The developer should use surrogate samples derived from the patient sample types intended for use with the automated system. As shown in Table 16, the preferred surrogate sample type is an unspiked pooled patient sample (A2), followed by individual or pooled patient samples spiked with the target analyte (B1 preferred over B2). The surrogate samples are spiked to an appropriately high analyte concentration to measure the effect of sample and/or analyte carryover. Before examination, surrogate samples should be processed with the same preexamination steps as patient specimens (eg, in a sample buffer). With proper justification, artificial analyte may be acceptable for use (E1 preferred over E2). When the analyte is an exogenous material (eg, for qualitative infectious diseases assays) and blank patient samples are needed but have insufficient volume (eg, negative swab samples), using a simulated artificial matrix (G, H, I) ensures that the samples are truly negative. In this case, using

simulated artificial matrix is appropriate because hundreds of tests are typically needed for the carryover study, and it would be difficult to ensure that such a large pool of patient swab samples is truly negative. Detection of low analyte concentrations in such a pool may be mistakenly counted as carryover events.

When feasible, only surrogate samples E2 or higher in the surrogate sample hierarchy should be used. The rationale is to maintain matrix integrity and not compromise the matrix's key characteristics (eg, viscosity, ionic strength), which could negatively affect the data and make it impossible to accurately assess carryover or cross-contamination. The developer may consider using artificial matrixes when the volume of individual patient samples is low and it is difficult to obtain the volume needed for a carryover study (eg, patient swab samples for a qualitative infectious diseases assay).

Special consideration is needed for sample types with inherent biological variability that may be difficult to mimic with a surrogate sample (eg, excessively mucoidal samples, sputum, hemolyzed blood samples).

Table 16. Carryover Study Hierarchy⁴

Carryover Sample	Sample Definition
A1	Patient sample (unspiked, individual)
A2	Pooled (unspiked, pool)
B1	Supplemented (biological spiked, individual)
B2	Pooled (biological spiked, pool)
E1	Simulated analyte (artificial spiked, individual)
E2	Simulated analyte (artificial spiked, pool)
G	Simulated matrix (unspiked, artificial)
H	Simulated matrix (biological spiked, artificial)
I	Simulated matrix (artificial spiked, artificial)

8.10 Reagent Stability

8.10.1 Objective

A stable reagent maintains its performance characteristics within acceptable limits over an established time period. The objective of a reagent stability study is to determine whether the key performance metrics of an IVD or other test reagent continue to meet predefined acceptance criteria within the established expiration date when stored and handled as required.

8.10.2 Principles for Surrogate Sample Use

Developers establish shelf-life and in-use stability for quantitative and qualitative reagents. The acceptability of using surrogate samples depends on test design. For example, for establishment of short-term on-board reagent stability, patient samples of sufficient volume that contain a reasonable range of analyte concentrations are preferable. However, regardless of reagent use or storage conditions, reagent stability studies are usually longer term. To establish the performance claims, the developer often needs to demonstrate that significant sample volumes are stable over the anticipated study period. Therefore, surrogate samples may be needed to obtain the necessary sample volume.

8.10.3 When to Use

Surrogate samples can be used when individual patient samples lack the volume and/or the stability necessary to evaluate the assay over the study period. Using surrogate samples increases sample volume, ensures the availability of reproducible samples with appropriate analyte concentrations, and reduces or eliminates performance variation.

8.10.4 How to Use

Surrogate samples should supplement, not replace, patient samples, covering concentrations that patient samples do not. For example, if a low-concentration sample (eg, below the biological reference interval) is unavailable, the developer may need to use a surrogate pool or artificial matrix. If patient samples are unstable over the study period, surrogate samples can be used to ensure that any stability shift is due to the reagent, not the sample. Testing patient and surrogate samples in the same study is acceptable.

Patient sample pools, controls, and other suitable materials of known analyte stability are appropriate surrogate sample types (refer to CLSI document EP25³²). For quantitative tests, surrogate samples may be neat, diluted, or supplemented to represent the MDLs, LLMI, and upper limit of the measuring interval (ULMI). For qualitative tests, surrogate samples may be neat, diluted, or supplemented to represent the MDL or cutoff.

The selection of surrogate samples for reagent stability studies depends on the test, study design, availability of patient samples, and ability to stabilize (eg, freeze) the sample (see Table 17). Depending on the situation, more than one surrogate sample type may be used. Rather than following a hierarchy, the developer should use these points to guide decisions:

- In order to support the use of surrogate samples in stability studies, the developer needs to evaluate their performance in other performance studies.
- The testing plan needs to describe the surrogate sample type, as well as its desired stability.
- When surrogate samples are used, recognition of the analyte can change over time. The testing plan should account for this possibility.
- The developer should include (when available) patient samples with concentrations comparable to that of one surrogate sample or at the MDL.
- Analyte concentrations that are best suited for monitoring or evaluating test performance should be included. Options include concentrations that are near the LLoD or the ULMI or that cover the AMI, an MDL, or the interval over which the signal (ie, physical parameter) is most sensitive to change. When patient samples are unavailable, the defined, reproducible properties of surrogate samples make them useful for assessing important concentrations.
- When the analyte is stable under the required storage conditions, individual patient samples should be pooled so that their concentrations reach the MDLs, the LLMI, and the ULMI.
- In some cases (eg, when the matrix is labile and unsuitable for long-term storage), samples should be spiked with analyte so that their concentrations reach the MDLs, the LLMI, and the ULMI.
- Commercially available calibrator and control materials (often in lyophilized form) intended for use in calibration and routine QC monitoring may be used to supplement patient samples or verify test performance. However, they should not be used alone (ie, to the exclusion of other surrogate sample types), except when using surrogate samples composed of patient samples is not feasible. In this case, the testing plan should include a clear scientific rationale for exclusive use of calibrators and controls as surrogate samples.

- If the analyte is unstable under standard long-term storage conditions (eg, conjugates, some proteins), biosimilar materials may be used (eg, taurine-conjugated bilirubin, tanned cells, identified epitopes [for immunoassays]).
- Because sample extract that contains analyte is frequently more stable than sample-like material, it may be preferred for monitoring reagent performance. For qualitative assays, the optimal analyte concentration for monitoring may vary, depending on technology and the test's intended use.
- Surrogate samples should be stored under stable conditions (eg, -70°C).
- For reagents with a long shelf-life (eg, > 2 years), requalifying the surrogate samples or preparing new ones may be necessary to ensure sample stability.

Table 17. Reagent Stability Study Hierarchy⁴

Reagent Stability Sample	Sample Definition
A1	Patient sample (unspiked, individual)
A2	Pooled (unspiked, pool)
B1	Supplemented (biological spiked, individual)
B2	Pooled (biological spiked, pool)
E1	Simulated analyte (artificial spiked, individual)
E2	Simulated analyte (artificial spiked, pool)
G	Simulated matrix (unspiked, artificial)
H	Simulated matrix (biological spiked, artificial)
I	Simulated matrix (artificial spiked, artificial)

Chapter 9

Conclusion

9 Conclusion

To facilitate scientific advancements, this guideline establishes a definition of “surrogate sample.” Adoption of this terminology by stakeholders is expected to improve the exchange of information and promote appropriate study design. EP39 describes a methodical approach for determining the suitability of using surrogate samples in different situations. It discusses the potential limitations of surrogate samples, as well as mechanisms to manage these limitations through appropriate sample construction and performance study design. Owing to the inherent difficulties of procuring patient samples, such as the rarity of certain sample types, low prevalence of some diseases, low sample volumes, and limited quantities at necessary concentrations, surrogate samples play an important role in diagnostic and laboratory science. Application of this guideline will simplify terminology for surrogate samples and expedite scientific developments.

Chapter 10

Supplemental Information

This chapter includes:

- References
- Appendixes
- The Quality Management System Approach
- Related CLSI Reference Materials

References

- 1 CLSI. *A Framework for Using CLSI Documents to Evaluate Clinical Laboratory Measurement Procedures*. 2nd ed. CLSI report EP19. Clinical and Laboratory Standards Institute; 2015.
- 2 Miller JM, Astles R, Baszler T, et al. Guidelines for safe work practices in human and animal medical diagnostic laboratories. Recommendations of a CDC-convened, Biosafety Blue Ribbon Panel. *MMWR Suppl*. 2012;61(1): 1-102.
- 3 CLSI. *Protection of Laboratory Workers From Occupationally Acquired Infections; Approved Guideline—Fourth Edition*. CLSI document M29-A4. Clinical and Laboratory Standards Institute; 2014.
- 4 Medical Device Innovation Consortium (MDIC) Clinical Diagnostics Surrogate Sample Use Working Group. *Surrogate Sample Framework: A Report of the Surrogate Sample Working Group of the Medical Device Innovation Consortium (MDIC)*. Medical Device Innovation Consortium; 2017. Accessed 13 September 2021.
<https://mdic.org/wp-content/uploads/2017/09/MDIC-Surrogate-Sample-Framework.pdf>
- 5 ISO. *In vitro diagnostic medical devices – Requirements for establishing metrological traceability of values assigned to calibrators, trueness control materials and human samples*. ISO 17511. International Organization for Standardization; 2020.
- 6 Bureau International des Poids et Mesures (BIPM). *International Vocabulary of Metrology – Basic and General Concepts and Associated Terms* (VIM, 3rd edition, JCGM 200:2012). Accessed 13 September 2021.
<https://www.bipm.org/en/publications/guides/>
- 7 ISO. *In vitro diagnostic medical devices – Measurement of quantities in samples of biological origin – Requirements for content and presentation of reference measurement procedures*. ISO 15193. International Organization for Standardization; 2009.
- 8 ISO. *Medical laboratories – Requirements for quality and competence*. ISO 15189. International Organization for Standardization; 2012.
- 9 ISO. *In vitro diagnostic medical devices – Information supplied by the manufacturer (labelling) – Part 1: Terms, definitions and general requirements*. ISO 18113-1. International Organization for Standardization; 2009.
- 10 ISO. *In vitro diagnostic medical devices – Measurement of quantities in biological samples – Metrological traceability of values for catalytic concentration of enzymes assigned calibrators and control materials*. ISO 18153. International Organization for Standardization; 2003.
- 11 ISO. *In vitro diagnostic medical devices – Measurement of quantities in samples of biological origin – Requirements for certified reference materials and the content of supporting documentation*. ISO 15194. International Organization for Standardization; 2009.
- 12 Miller WG. Specimen materials, target values and commutability for external quality assessment (proficiency testing) schemes. *Clin Chim Acta*. 2003;327(1-2):25-37. doi:10.1016/s0009-8981(02)00370-4

- ¹³ ISO. *Medical devices – Application of risk management to medical devices*. ISO 14971. International Organization for Standardization; 2019.
- ¹⁴ ISO. *Medical laboratories – Application of risk management to medical laboratories*. ISO 22367. International Organization for Standardization; 2020.
- ¹⁵ ISO. *Medical devices – Guidance on the application of ISO 14971*. ISO/TR 24971. International Organization for Standardization; 2020.
- ¹⁶ CLSI. *Laboratory Quality Control Based on Risk Management; Approved Guideline*. CLSI document EP23-A™. Clinical and Laboratory Standards Institute; 2011.
- ¹⁷ CLSI. *Risk Management Techniques to Identify and Control Laboratory Error Sources; Approved Guideline—Second Edition*. CLSI document EP18-A2. Clinical and Laboratory Standards Institute; 2009.
- ¹⁸ CLSI. *User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline—Second Edition*. CLSI document EP12-A2. Clinical and Laboratory Standards Institute; 2008.
- ¹⁹ CLSI. *Molecular Diagnostic Methods for Infectious Diseases*. 3rd ed. CLSI report MM03. Clinical and Laboratory Standards Institute; 2015.
- ²⁰ CLSI. *Evaluation of Linearity of Quantitative Measurement Procedures*. 2nd ed. CLSI guideline EP06. Clinical and Laboratory Standards Institute; 2020.
- ²¹ CLSI. *Evaluation of Detection Capability for Clinical Laboratory Measurement Procedures; Approved Guideline—Second Edition*. CLSI document EP17-A2. Clinical and Laboratory Standards Institute; 2012.
- ²² CLSI. *Measurement Procedure Comparison and Bias Estimation Using Patient Samples*. 3rd ed. CLSI guideline EP09c. Clinical and Laboratory Standards Institute; 2018.
- ²³ CLSI. *Quantitative Molecular Methods for Infectious Diseases; Approved Guideline—Second Edition*. CLSI document MM06-A2. Clinical and Laboratory Standards Institute; 2010.
- ²⁴ CLSI. *Assessment of Equivalence or Suitability of Specimen Types for Medical Laboratory Measurement Procedures*. 1st ed. CLSI guideline EP35. Clinical and Laboratory Standards Institute; 2019.
- ²⁵ CLSI. *Evaluation of Commutability of Processed Samples; Approved Guideline—Third Edition*. CLSI document EP14-A3. Clinical and Laboratory Standards Institute; 2014.
- ²⁶ CLSI. *Evaluation of Precision of Quantitative Measurement Procedures; Approved Guideline—Third Edition*. CLSI document EP05-A3. Clinical and Laboratory Standards Institute; 2014.
- ²⁷ CLSI. *Interference Testing in Clinical Chemistry*. 3rd ed. CLSI guideline EP07. Clinical and Laboratory Standards Institute; 2018.
-

- ²⁸ CLSI. *Establishing and Verifying an Extended Measuring Interval Through Specimen Dilution and Spiking*. 1st ed. CLSI guideline EP34. Clinical and Laboratory Standards Institute; 2018.
- ²⁹ CLSI. *Preparation and Validation of Commutable Frozen Human Serum Pools as Secondary Reference Materials for Cholesterol Measurement Procedures; Approved Guideline*. CLSI document C37-A. Clinical and Laboratory Standards Institute; 1999.
- ³⁰ ISO. *Biological evaluation of medical devices – Part 12: Sample preparation and reference materials*. ISO 10993-12. International Organization for Standardization; 2021.
- ³¹ CLSI. *Analysis of Body Fluids in Clinical Chemistry*. 2nd ed. CLSI guideline C49. Clinical and Laboratory Standards Institute; 2018.
- ³² CLSI. *Evaluation of Stability of In Vitro Diagnostic Reagents; Approved Guideline*. CLSI document EP25-A. Clinical and Laboratory Standards Institute; 2009.
- ³³ ISO. *In vitro diagnostic medical devices – Evaluation of stability of in vitro diagnostic reagents*. ISO 23640. International Organization for Standardization; 2011.
- ³⁴ CLSI. *Characterization and Qualification of Commutable Reference Materials for Laboratory Medicine; Approved Guideline*. CLSI document EP30-A. Clinical and Laboratory Standards Institute; 2010.
- ³⁵ Miller WG, Schimmel H, Rej R, et al. IFCC Working Group Recommendations for Assessing Commutability Part 1: General Experimental Design. *Clin Chem*. 2018;64(3):447-454. doi:10.1373/clinchem.2017.277525
- ³⁶ Nilsson G, Budd JR, Greenberg N, et al. IFCC Working Group Recommendations for Assessing Commutability Part 2: Using the Difference in Bias between a Reference Material and Clinical Samples. *Clin Chem*. 2018;64(3):455-464. doi:10.1373/clinchem.2017.277541
- ³⁷ Budd JR, Weykamp C, Rej R, et al. IFCC Working Group Recommendations for Assessing Commutability Part 3: Using the Calibration Effectiveness of a Reference Material. *Clin Chem*. 2018;64(3):465-474. doi:10.1373/clinchem.2017.277558
- ³⁸ CLSI. *Validation and Verification of Multiplex Nucleic Acid Assays*. 2nd ed. CLSI guideline MM17. Clinical and Laboratory Standards Institute; 2018.
- ³⁹ Hassoun M, Royall PG, Parry M, Harvey RD, Forbes B. Design and development of a biorelevant simulated human lung fluid. *J Drug Deliv Sci Technol*. 2018;47:485-491. doi:10.1016/j.jddst.2018.08.006
- ⁴⁰ ISO. *Textiles – Tests for colour fastness – Part E04: Colour fastness to perspiration*. ISO 105-E04. International Organization for Standardization; 2013.
- ⁴¹ Colorfastness to perspiration: AATTCC Test Method 15 - 2002. Accessed 13 September 2021. <https://studylib.net/doc/11029665/colorfastness-to-perspiration-aattcc-test-method-15---2002>
- ⁴² Marques MRC, Loebenberg R, Almukainzi M. Simulated biological fluids with possible application in dissolution testing. *Dissolut Technol*. 2011;18(3):15-28. doi:10.14227/dt180311p15
-

- ⁴³ CLSI. *Supplemental Tables for Interference Testing in Clinical Chemistry*. 1st ed. CLSI supplement EP37. Clinical and Laboratory Standards Institute; 2018.
- ⁴⁴ Broughton PM. Carry-over in automatic analysers. *J Automat Chem*. 1984;6(2):94-95. doi:10.1155/s1463924684000201
- ⁴⁵ Haeckel R. Proposals for the description and measurement of carry-over effects in clinical chemistry (Recommendations 1991). *Pure Appl Chem*. 1991;63(2):301-306. doi:10.1351/pac199163020301

Additional Resources

About the Specimen Resource Locator. National Cancer Institute Specimen Resource Locator. Accessed 13 September 2021.
<https://specimens.cancer.gov/tissue/default.htm>

CDC & FDA Antibiotic Resistance (AR) Isolate Bank. CDC. Updated 9 March 2020. Accessed 13 September 2021.
<https://www.cdc.gov/drugresistance/resistance-bank/index.html>

CEN-CENELEC. Accessed 13 September 2021. <https://www.cencenelec.eu/>

Laboratory quality assurance and standardization programs. CDC. Updated 6 July 2017. Accessed 13 September 2021.
<https://www.cdc.gov/labstandards/overview.html>

NIST TN 1297: Appendix D4. Measurand defined by the measurement method; characterization of test methods; simple calibration. NIST. Published 6 November 2015. Updated 25 November 2019. Accessed 13 September 2021.

<https://www.nist.gov/pml/nist-technical-note-1297/nist-tn-1297-appendix-d4-measurand-defined-measurement-method>

Standard Reference Materials. NIST. Accessed 13 September 2021. <https://www.nist.gov/srm>

Tuberculosis Frozen Trial Repository. FDA. Accessed 13 September 2021. https://www.accessdata.fda.gov/scripts/fdatrak/view/track_project.cfm?program=oc-chief-scientist&id=OCS-ORSI-Tuberculosis-Frozen-Trial-Repository

WHO international standards. National Institute for Biological Standards and Control. Accessed 13 September 2021.
https://www.nibsc.org/products/brm_product_catalogue/who_standards.aspx

WHO/TDR TB Specimen Bank. World Health Organization; Special Programme for Research and Training in Tropical Diseases. Accessed 13 September 2021.

http://www.stoptb.org/wg/new_diagnostics/assets/documents/Information_SpecimenBank_V2.pdf

Appendix A. Surrogate Sample Descriptions and Examples

Appendix A is modified from the Medical Device Innovation Consortium (MDIC) Clinical Diagnostics Surrogate Sample Use Working Group. *Surrogate Sample Framework: A Report of the Surrogate Sample Working Group of the Medical Device Innovation Consortium (MDIC)*. Medical Device Innovation Consortium; 2017, with permission.

Alphanumeric Designation	Sample Description	Example(s)
A1	Patient sample; includes prospective, retrospective, and archived patient specimens, plus patient specimens prepared with assay process steps (eg, extraction)	<ul style="list-style-type: none">Hyperlipidemic plasma specimen, plasma specimen from patient with Crigler-Najjar syndrome (eg, conjugated hyperbilirubinemia), high-HDL cholesterol plasma specimen (with extraction needed)Patient specimen (eg, plasma, serum, whole blood, FFPE tissue, swab, or sample in collection media)Patient stool specimen processed through sample preparation steps of assayPatient sputum specimen processed through sample preparation steps of assay
A2	Pooled individual patient specimens	<ul style="list-style-type: none">Combined plasma specimens with high hormone concentrations (eg, plasma from patients with Cushing syndrome, plasma from pregnant women, plasma with elevated enzyme activity)Pool of patient specimens (eg, plasma, serum, whole blood, or sample in collection media)A fixed number (eg, 3) of known assay-positive patient stool specimens combinedA fixed number (eg, 3) of known assay-positive patient sputum specimens combinedA fixed number (eg, 3) of known negative patient serum specimens combined
B1	Individual patient specimen supplemented with biological analyte	Material from a patient specimen with high analyte concentration spiked into another patient specimen (eg, plasma, serum, whole blood, FFPE tissue, swab, or sample in collection media)
B2	Pooled patient specimens supplemented with biological analyte	<ul style="list-style-type: none">Material from a pool of patient specimens spiked into a different pool of patient specimens (eg, plasma, serum, whole blood, FFPE tissue, swab, or sample in collection media)A fixed number (eg, 2) of known negative patient stool specimens combined and spiked with a known assay-positive patient specimenA fixed number (eg, 2) of known negative patient sputum specimens combined and spiked with a known assay-positive patient specimen

Appendix A. (Continued)

Alphanumeric Designation	Sample Description	Example(s)
C1	Individual patient specimen that has been altered	<ul style="list-style-type: none"> Specimen used for hormone analysis that has undergone heat treatment or stripping of analyte when absence or low concentration of hormone has clinical implications Patient source (from apheresis) plasma specimen Known assay-positive patient sputum specimen diluted with assay sample preparation buffer
C2	Pooled patient specimens that have been altered	<ul style="list-style-type: none"> Serum from 3 different patients combined and heat treated Pool of patient source (from apheresis) plasma specimens A fixed number (eg, 3) of known assay-positive patient sputum specimens combined and diluted with assay sample preparation buffer
D1	Individual patient specimen that has been altered and supplemented with biological analyte	Patient source (from apheresis) plasma specimen spiked into another patient source (from apheresis) plasma specimen
D2	Pooled patient specimens that have been altered and supplemented with biological analyte	Pool of patient source (from apheresis) plasma specimens spiked into a different pool of patient source (from apheresis) plasma specimens
E1	Individual patient specimen supplemented with simulated analyte	<ul style="list-style-type: none"> Plasmid, armored RNA, viral culture, bacterial culture, cell-line DNA, or WHO or other standard spiked into an individual patient specimen (eg, plasma, serum, whole blood, FFPE tissue, swab, or sample in collection media) Animal-sourced homolog spiked into an individual patient specimen (eg, taurine-conjugated bilirubin, salt version of drug used to enhance solubility into individual plasma or urine)
E2	Pooled patient specimens supplemented with simulated analyte	Plasmid, armored RNA, viral culture, bacterial culture, cell-line DNA, or WHO or other standard spiked into a pool of patient specimens (eg, plasma, serum, whole blood, FFPE tissue, swab, or sample in collection media)

Appendix A. (Continued)

Alphanumeric Designation	Sample Description	Example(s)
F1	Individual patient specimen that has been altered and supplemented with simulated analyte	<ul style="list-style-type: none"> • Plasmid, armored RNA, viral culture, or WHO or other standard spiked into a patient source (from apheresis) plasma specimen • Organic stock solution of drug spiked into an individual patient specimen, with matrix altered to a defined degree • Isolated antigen or primer added to extraction media for certain methods (eg, GC, LC, LC-MS, NAD) performed on an individual patient specimen
F2	Pooled patient specimens that have been altered and supplemented with simulated analyte	Plasmid, armored RNA, viral culture, or WHO or other standard spiked into a pool of patient source (from apheresis) plasma specimens
G	Artificial matrix	Diluent
H	Artificial matrix supplemented with biological analyte	Artificial vaginal sample matrix spiked with known high-analyte vaginal swab sample
I	Artificial matrix supplemented with simulated analyte	Plasmid, armored RNA, or cell-line DNA spiked into a sample composed of diluted whole blood

Abbreviations: DNA, deoxyribonucleic acid; FFPE, formalin-fixed, paraffin-embedded; GC, gas chromatography; HDL, high-density lipoprotein; LC, liquid chromatography; LC-MS, liquid chromatography-mass spectrometry; NAD, nicotinamide adenine dinucleotide; NALC-NaOH, N-acetyl-L-cysteine-sodium hydroxide; RNA, ribonucleic acid; WHO, World Health Organization.

Appendix B. Example Chemical Compositions for Artificial Matrixes

Tables B1 to B3 provide example chemical compositions for artificial matrixes. They are meant as a reference point and do not represent all possible formulations.

Table B1. BAL, CSF, Intestinal Fluid, and Saliva

BAL ¹	BAL Concentration	CSF ^{2,3}	CSF Concentration, g/L	Intestinal Fluid ⁴	Intestinal Fluid Concentration	Saliva ⁵	Saliva Concentration, g/L
DPPC	4.8 mg/mL	NaCl ⁻	7.592	KH ₂ PO ₄	68.05 g	NaCl ⁻	8.0
DPPG	0.5 mg/mL	KCl	0.146	NaOH	8.96 g	KH ₂ PO ₄	0.19
Cholesterol	0.1 mg/mL	MgCl ₂ • 6H ₂ O	0.24	Deionized water	to 10.0 L	Na ₂ HPO ₄	2.38
Albumin	8.8 mg/mL	CaCl ₂ • 2H ₂ O	0.174				
IgG	2.6 mg/mL	KH ₂ PO ₄	0.154				
Transferrin	1.5 mg/mL	Sodium lactate	1.69				
Ascorbate	140 µM	Glucose	0.719				
Urate	95 µM	NaHCO ₃	0.924				
Glutathione	170 µM						

Abbreviations: BAL, bronchoalveolar lavage; CSF, cerebrospinal fluid; DPPC, dipalmitoylphosphatidylcholine; DPPG, dipalmitoylphosphatidylglycerol; IgG, immunoglobulin G.

Table B2. Semen, Sputum, Sweat, and Synovial Fluid

Semen ⁶	Semen Concentration, g/L	Sputum ⁷	Sweat ⁸	Sweat Concentration, g/L	Synovial Fluid ^{9,a}	Synovial Fluid Concentration, g/L
Glucose ⁻	1.02		NaCl ⁻	0.5	Na ⁺	153.1
Sodium citrate	8.13	1 emulsified egg to 1 L 1% (w/v) aqueous methylcellulose prepared with sterile water	Lactic acid	0.1	K ⁺ buffer	4.2
KOH	0.881		Urea	0.1	Cl	139.6
KCl ⁻	0.908				Phosphate buffer	9.6
Fructose	2.72					
Lactic acid	0.62					
Urea	0.42					
BSA	50.4					
Na ₂ HPO ₄ • H ₂ O	0.0927					
Na ₂ HPO ₄	0.858					

Abbreviation: BSA, bovine serum albumin.

^a Cl contains 0.3% weight percent hyaluronic acid.

Appendix B. (Continued)

Table B3. Tears, Urine, Vaginal Fluid, and Serum

Tears ¹⁰	Tears Concentration, g/L	Urine ¹¹	Urine Concentration	Vaginal Fluid ¹²	Vaginal Fluid Concentration, g/L	Serum ¹³	Serum Concentration
Lysozyme	1.9	NaCl ⁻	14.1 g/L	Glucose	5.0	PBS	150 mM
Lactoferrin	1.8	KCl ⁻	2.8 g/L	Glycerol	0.16	BSA	40-60 g/L (fit adjusted for purpose)
α -acid glycoprotein	0.5	Urea	17.3 g/L	Lactic acid	2.0		
Albumin	0.2	25% ammonia prepared in water	0.19%	Acetic acid	1.0		
Mucin	0.15	CaCl ₂	0.6 g/L	Albumin	0.018		
γ -globulins	0.1	MgSO ₄	0.43 g/L	Urea	0.40		
NaCl ⁻	6.626	HCl ⁻	0.02 M	NaCl ⁻	3.51		
KCl ⁻	1.716			KOH	1.40		
NaHCO ₃	1.376			Ca(OH) ₂	0.222		
Lactic acid	0.27			pH	4.2		
CaCl ₂	0.147						
NaH ₂ PO ₄ • H ₂ O	0.1						
Lipids	0.0798						

Abbreviations: BSA, bovine serum albumin; PBS, phosphate-buffered saline; pH, negative logarithm of hydrogen ion concentration.

References for Appendix B

- 1 Kumar A, Terakosolphan W, Hassoun M, et al. A biocompatible synthetic lung fluid based on human respiratory tract lining fluid composition. *Pharm Res.* 2017;34(12):2454-2465. doi: 10.1007/s11095-017-2169-4
- 2 Oka K, Yamamoto M, Nonaka T, Tomonaga M. The significance of artificial cerebrospinal fluid as perfusate and endoneurosurgery. *Neurosurgery.* 1996;38(4):733-736. doi: 10.1227/00006123-199604000-00019
- 3 Spector R, Snodgrass SR, Johanson CE. A balanced view of the cerebrospinal fluid composition and functions: focus on adult humans. *Exp Neurol.* 2015;273:57-68. doi: 10.1016/j.expneurol.2015.07.027
- 4 Stippler E, Kopp S, Dressman JB. Comparison of US Pharmacopeia simulated intestinal fluid TS (without pancreatin) and phosphate standard buffer pH 6.8, TS of the International Pharmacopoeia with respect to their use in *in vitro* dissolution testing. *Dissolut Technol.* 2004;11(2):6-10. doi: 10.14227/dt110204p6
- 5 Mashru RC, Sutariya VB, Sankalia MG, Parikh PP. Development and evaluation of fast-dissolving film of salbutamol sulphate. *Drug Dev Ind Pharm.* 2005;31(1):25-34. doi: 10.1081/ddc-43947
- 6 Owen DH, Katz DF. A review of the physical and chemical properties of human semen and the formulation of a semen simulant. *J Androl.* 2005;26(4):459-469. doi: 10.2164/jandrol.04104

Appendix B. (Continued)

- 7 Aber VR, Allen BW, Mitchison DA, Ayuma P, Edwards EA, Keyes AB. Quality control in tuberculosis bacteriology. 1. Laboratory studies on isolated positive cultures and the efficiency of direct smear examination. *Tubercle*. 1980;61(3):123-133. doi: 10.1016/0041-3879(80)90001-x
- 8 BSI. *Reference test method for release of nickel from all post assemblies which are inserted into pierced parts of the human body and articles intended to come into direct and prolonged contact with the skin*. BS EN 1811:2011+A1:2015. British Standards Institution; 2015.
- 9 Conzone SD, Brown RF, Day DE, Ehrhardt GJ. *In vitro* and *in vivo* dissolution behavior of a dysprosium lithium borate glass designed for the radiation synovectomy treatment of rheumatoid arthritis. *J Biomed Mater Res*. 2002;60(2):260-268. doi: 10.1002/jbm.10047
- 10 Mirejovsky D, Patel AS, Rodriguez DD, Hunt TJ. Lipid adsorption onto hydrogel contact lens materials: advantages of Nile red over oil red O in visualization of lipids. *Optom Vis Sci*. 1991;68(11):858-864. doi: 10.1097/00006324-199111000-00005
- 11 Tanaka T, Hayashi Y. Determination of silicon calcium, magnesium and phosphorus in urine using inductively-coupled plasma emission spectrometry and a matrix-matching technique. *Clin Chem Acta*. 1986;156(1):109-113. doi: 10.1016/0009-8981(86)90186-5
- 12 Owen DH, Katz DF. A vaginal fluid simulant. *Contraception*. 1999;59(2):91-95. doi: 10.1016/s0010-7824(99)00010-4
- 13 Van de Merbel NC. Quantitative determination of endogenous compounds in biological samples using chromatographic techniques. *Trends Analytic Chem*. 2008;27(10):924-933. doi: 10.1016/j.trac.2008.09.002

Appendix C. *Chlamydia trachomatis*: Precision/Reproducibility

Appendix C contains excerpts from Medical Device Innovation Consortium (MDIC). *Surrogate Sample Framework Harmonized Education Case Study: Chlamydia trachomatis - Precision/Reproducibility*. Arlington, VA: Medical Device Innovation Consortium; 2019, with permission, and has been maintained verbatim from the source material.

NOTE: This case study is written in a conversational tone.

C1 Abbreviations and Acronyms

CT *Chlamydia trachomatis*

C2 Introduction

This is the first of five case studies intended to demonstrate how to apply the Surrogate Sample Framework. In this case study, we will review precision, including reproducibility, for *Chlamydia trachomatis* tests. Please note that the types of specimens you consider are important. In this case study, we will first consider female and male fresh urine, followed by a discussion using the Surrogate Sample Framework for vaginal or endocervical swabs.

C3 Principles

There are seven principles for this case study.

C3.1 Principle 1 - Identify the Objective for Using Surrogate Samples for This Study

First, let's review the precision study and explore the challenges that are often encountered during this performance testing. In a precision study for a qualitative test, we need to have samples with analyte concentrations close to the cutoff, because precision around the cutoff is an important characteristic. However, patient samples with concentrations around the cutoff may be difficult to find. While conducting a precision study, we may need to test the sample many times under different testing conditions, so we need to have many replicates from the same sample. This can present challenges with sample volume.

Precision samples should be tested over five or more days (for example, a within-laboratory precision study may cover 20 days), so sample stability over the duration of the entire study is an important consideration. If a reproducibility study requires sample testing at different sites, extending the sample stability considerations to include transportation is important. Due to these potential stability challenges, using surrogate samples in such situations has clear benefits if native samples do not have the necessary stability.

Appendix C. (Continued)

Another advantage of using a surrogate sample is that we can create a sample with an analyte concentration close to the cutoff, should it be difficult to do so with native patient samples. Even though this is a **qualitative** test, we need to test samples over the dynamic range of the test, which spans low, mid, and high measurement levels of analyte, just as we would with a quantitative test. Another potential benefit, pooling multiple samples will reduce the chance of introducing bias from an atypical individual sample.

C3.2 Principle 2 - Identify Critical Factors Related to the Performance Study Type

We briefly reviewed the challenges associated with precision studies, including reproducibility studies. To recap here, it is important to have a large volume of sample. This sample should be homogenous and remain stable for the duration of the precision/reproducibility studies. For example, performing the within-laboratory precision study noted in the previous section would require a sample that is stable for at least 20 days, the duration of the study.

C3.3 Principle 3 - Consider Factors Based on Whether the Test Is Qualitative, Quantitative, or Semi-Quantitative

Another important principle is to consider whether the test being assessed is qualitative, quantitative, or semi-quantitative. In this case study, the chlamydia IVD test is a qualitative, binary test. Results of the test are binary: positive or negative, yes or no, detected or not detected. Due to the binary nature of the results, it is important to have samples near the cutoff.

First, we need to have a negative sample. Negative samples can also be called “true negatives”. This sample would have zero concentration of analyte; in other words, we should know with certainty that analyte is absent. As mentioned previously, we also need a sample that is close to the cutoff, specifically a positive C95 sample, sometimes referred to as a “low-positive sample”. “Positive C95” means the amount of analyte in the sample is just above the cutoff, and when you repeatedly test this sample, 95% of the time it will have positive results and 5% of the time it will have negative results. In addition, we need to have a “medium-positive” sample with a value that is relatively far from the cutoff and has positive results 100% of the time. Finally, we need to have a “high-positive” sample with a value that is very far from the cutoff and always has a positive result. The high-positive sample should ideally be close to the upper limit of the detection range for the assay and is used to assess signal variability at the upper limit of the assay range.

To recap, we need the following precision samples:

1. Negative sample
2. C95 or low-positive sample
3. Medium-positive sample
4. High-positive sample

C3.4 Principle 4 - Consider the Effect of Sample Processing Such as Extraction Method, If Applicable

Patient samples for chlamydia assays typically undergo a processing step. Therefore, the surrogate sample should undergo that same processing step, because the surrogate should be treated as a patient sample to the greatest extent possible. If lysis of the *Chlamydia*

Appendix C. (Continued)

trachomatis (CT) organism is part of the processing step, then an intact organism of the most prevalent type of CT organism should be used. As there are different types of CT organism, consider whether all types or only the most prevalent type will be evaluated in the precision study. The most prevalent type should always be evaluated, and less prevalent types of the CT organism may need to be evaluated, depending on the goal of the study.

When preparing your surrogate sample, it is important to pay attention to how the sample will be transported to the laboratory where it will be tested. As with processing steps, the surrogate sample should be treated the same as a patient sample. We have two specimen types in this case study, a spiked collection swab and a urine sample, both of which are typically placed in a transport tube. Note that if testing is being performed at multiple sites, freezing of the aliquots should be considered to ensure samples are stable to transport. A spiked swab surrogate should be placed in the tube with transport media, and a urine sample might sometimes require a preservative. If interference studies showed that your technology is gender agnostic, either male urine or female urine can be included in the precision study. Otherwise, you should consider using a pool of male-only or female-only urine to minimize artificial effects in a precision study.

C3.5 Principle 5 - Evaluate Property Characteristics of the Matrix

We are now going to consider the set of principles in which we evaluate the matrix and the analyte. First, we evaluate the property characteristics of the matrix. For urine, we need a negative sample. It may be difficult to find a true-negative sample, because tests with high sensitivity can still have false-negative results. Even if the sample was tested and it has a negative result, we cannot exclude that this is a false-negative result (that is, that the amount of analyte was below the limit of detection of the test). A logical assumption would be that one can test the sample with multiple tests and if all results were negative, the sample can be considered a true negative. In the process of testing for a negative matrix, we could likely use up most of the sample and limit the volume available for the precision study. The more “true-negative” samples that need to be pooled, the greater the possibility that at least one of the samples was a false negative, meaning that the pool of these samples would then have some amount of analyte. Therefore, it is preferable to use an artificial matrix that simulates urine. This would likely be a buffer that contains cells and other substances that are present in urine.

Continuing with additional matrix considerations, human cells should be added to the matrix because they are found in a typical patient sample (and remember, we want our surrogates to mimic patient samples). There are other substances, such as blood from a heavy menses, that are atypical. Because we want to mimic representative patient samples, blood and other atypical substances should not be added to the matrix. We are interested in the precision evaluation for the typical patient, not the atypical patient.

C3.6 Principle 6 - Evaluate Property Characteristics of the Analyte

The analyte should mimic the natural state of the CT organism and be the most prevalent type of CT organism (though there may be a need for less prevalent types, depending on the test). The intact CT organism is preferred to cell lysate because lysate does not represent the natural state of the CT organism.

Appendix C. (Continued)

C3.7 Principle 7 - Assess Analyte and Matrix Combinations of Surrogate Samples

We have evaluated the properties of the analyte and the matrix separately. Now we must also assess the properties of the surrogate sample once we combine the analyte and matrix. We need to assess whether there are artificial effects when one combines some particular type of the analyte and some particular type of matrix. For example, there are some assays where an artificial analyte in a patient-derived matrix does not mimic patient samples, but using an artificial matrix with patient-derived analyte may only produce negligible effects. Surrogate sample selection requires a balance between the matrix type and analyte type, choosing the type of surrogate sample that most closely mimics the effects of a patient sample.

C4 Study-Specific Hierarchy: Urine Specimen

In the second half of this case study, we will navigate the study-specific hierarchy. We will walk through the different surrogate sample types and scenarios, eliminating choices as we go, to arrive at the most appropriate surrogate sample for this study. We will start with a urine specimen. This case study relates to a qualitative test, so we must have the following samples:

- **Negative**, denoted as N
- **Low positive**, denoted as LP, in which the concentration is close to the cutoff
- **Medium positive**, denoted as MP, in which the concentration is 100% positive at a mid-range point in the assay
- **High positive**, denoted as HP, in which the concentration is nearly at the upper limit of the range of the assay

Appendix C. (Continued)

Case Study 1

Hierarchy: CT - Urine Specimen

Precision
A1
A2
B1
B2
E1
E2
C1
C2
D1
D2
F1
F2
H
I
G

A1

Analyte: Unspiked, Individual

Matrix: Biological, Unaltered

Example
Urine sample from one patient, male or female

Challenges

- May have insufficient volume to perform a large study
- Longer lead time to collect sufficient samples for the study

Outcome

Go to the next level 

Figure C1. Analyte: Unspiked, Individual; Matrix: Unaltered, Biological

After we discuss sample composition, we'll address challenges associated with that sample. For example, the challenges associated with this A1 sample are volume limitations and longer lead time to collect samples. Then, based on the challenges we identify, we'll arrive at an outcome, which will either be to use the current tier of surrogate sample or go on to the next surrogate sample tier. Here, this A1 sample is very difficult to collect (for example, due to volume requirements), so we'll go to the next tier level.

Appendix C. (Continued)

Case Study 1

Hierarchy: CT - Urine Specimen

Precision
A1
A2
B1
B2
E1
E2
C1
C2
D1
D2
F1
F2
H
I
G

MP, HP

A2

Analyte:
Unspiked, Pooled

Matrix:
Biological, Unaltered

Example
Pooled urine from male
and/or female samples

Challenges

- N - Pool may not be true negative
- LP - With regard to low-positive sample, unable to target concentrations close to the cutoff
- MP/HP - With regard to medium or high positive, A2 may be appropriate
- All - This surrogate sample may be difficult to evaluate, such for a multiplex assay

Outcome

Appropriate for medium positive
and high positive; however, may
need to be spiked to reach
concentration level

Go to the
next level

Figure C2. Analyte: Unspiked, Pooled; Matrix: Unaltered, Biological

The analyte is unspiked, and the sample is one that is pooled from more than one patient. The matrix is biological and unaltered. An example of this A2 surrogate is a urine sample from a pool of single-gendered patients, either all female or all male. There are several challenges to consider for this sample:

- The negative sample pool might not be a true negative. Negative samples can have some amount of analyte that should also be considered in determining the proportions for pooling.
- For a low-positive sample, pooling can make it difficult to achieve a concentration that is close to the cutoff. For example, when we pool two samples, we need to understand what concentrations they have. Because the test is qualitative, there are signals for each of these samples, but the signals do not have a linear relationship with regard to concentration.
- A sample close to the cutoff for one qualitative test may not be close to the cutoff for another qualitative test (eg, a reference assay to establish the concentration).
- For medium- and high-positive samples, pooling of patient samples should be considered, if possible.
- Regarding high-positive samples, sometimes it is difficult to find samples that are close to the upper limit of the range. In this situation, we need to consider spiking the pool of patient samples in order to have a concentration close to the upper limit.

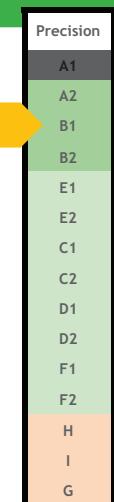
To summarize, we see that the A2 surrogate may be appropriate for medium and high positives. We have made a note by placing MP and HP to the right of A2 in the hierarchy on

Appendix C. (Continued)

the left side of the slide. Regarding the LP and the negative, A2 presents a lot of challenges. Therefore, we need to go to the next level.

Case Study 1

Hierarchy: CT - Urine Specimen



B1

Analyte:
Spiked -
Biological,
Individual

Matrix:
Biological,
Unaltered

Example

Material from one patient sample with high analyte concentration spiked into another patient sample. Urine with high CT concentration spiked into negative urine and varying dilutions.

A sample spiked with biological CT organisms reproduced by cell culture is considered a biological analyte.

Challenges

- All - May have insufficient volume to perform a large study
- All - Need to spike an amount that is close to the cutoff, what is the exact amount of high analyte concentration may be difficult to determine

Outcome

Go to the next level

Figure C3. Analyte: Spiked - Biological, Individual; Matrix: Unaltered, Biological

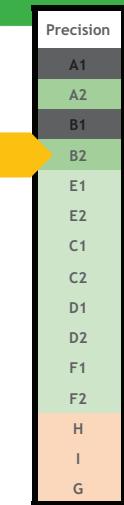
B1 means that the sample is from an individual patient and spiked with a biological analyte. The matrix of the patient sample is biological and unaltered. After spiking a biological analyte, the matrix should still be close to the patient matrix. An example of this B1 surrogate is material from a patient sample with a high CT concentration spiked into another patient sample. Another example is spiking negative urine with high-CT concentration urine and creating a dilution series with the negative urine sample. Note that we consider a sample spiked with biological CT organism reproduced by cell culture to be a “biological analyte”.

There are two challenges associated with using B1 for the surrogate samples needed in this study. Remember that B1 is an individual patient sample. Therefore, there are volume limitations, given a precision study may require many replicates. Also, it may be difficult to determine the appropriate dilution (proportion) of analyte to spike in order to achieve a concentration close to the cutoff.

Appendix C. (Continued)

Case Study 1

Hierarchy: CT - Urine Specimen



B2

Analyte:
Spiked -
Biological,
Pooled

Matrix:
Biological,
Unaltered

Example
Pool of known assay-negative patient samples combined and spiked with a known assay-positive patient sample.

A sample spiked with biological CT organisms reproduced by cell culture is considered a biological analyte.

Challenges

- N - Pool may not be true negative
- LP & HP - Volume of high concentration may be difficult to find
- LP - Need to spike an amount that is close to the cutoff, what is the exact amount of high analyte concentration may be difficult to determine

Outcome

[Go to the next level](#)

Figure C4. Analyte: Spiked - Biological, Pooled; Matrix: Unaltered, Biological

In B2, the sample is pooled from more than one patient, and it has a biological spike. The matrix is biological and unaltered, though it is a pool of a few patient samples. An example of this B2 surrogate is a pool of known negative patient samples spiked with urine from a patient with high CT values or spiked with CT organism reproduced by cell culture.

We encounter similar challenges for B2 as we have for the previous options:

- The negative sample might not be a true negative.
- For low- and high-positive samples, there may be volume limitations.
- For a low-positive sample, it may be difficult to determine the appropriate dilution to spike in to achieve a concentration close to the cutoff.

Appendix C. (Continued)

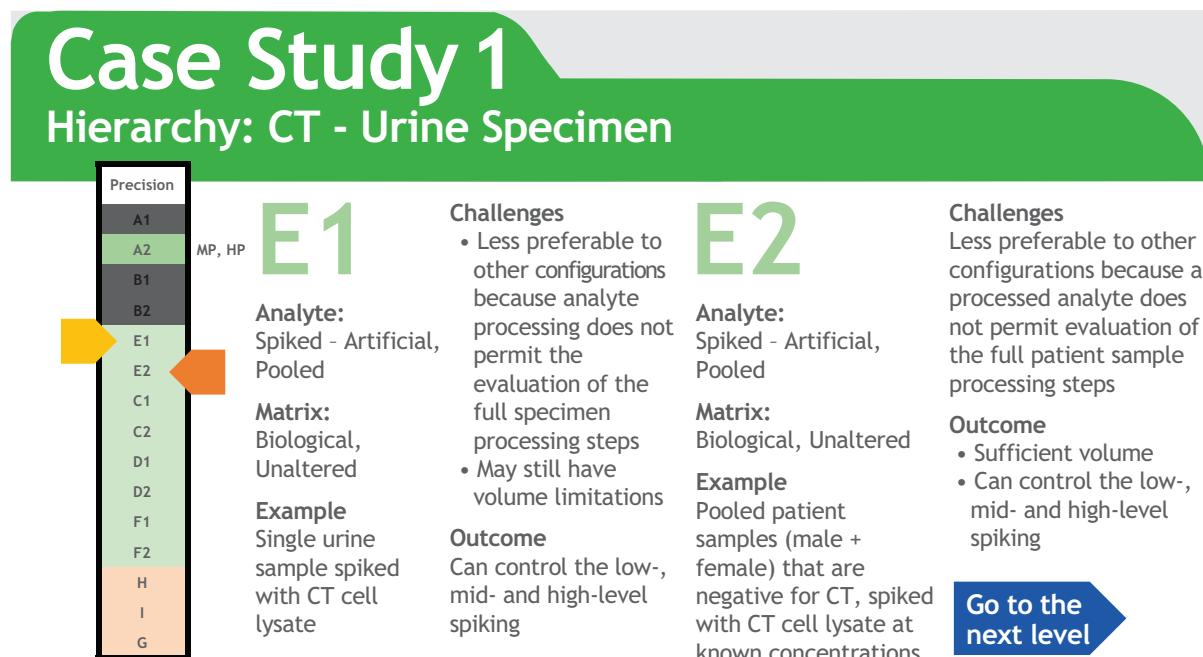


Figure C5. Analyte: Spiked - Artificial, Individual or Pooled; Matrix: Unaltered, Biological

In E1, the sample is from an individual patient and spiked with an artificial analyte.

The matrix is biological and unaltered. An example of an E1 surrogate sample is a single patient sample spiked with CT cell lysate. Note that for this type of surrogate sample, the spike is CT cell lysate, not CT organism reproduced by cell culture, and is therefore considered artificial. Using lysate is not preferred, because the lysis step does not mimic how patient samples are processed. Remember from our general principles that we want our surrogate sample preparation to mimic that of patient samples. Volume limitations may also be an issue for E1 surrogates, because we spike an individual sample.

Let's move to E2, where the sample is pooled from more than one patient and spiked with an artificial analyte. The matrix is biological and unaltered. An example of an E2 surrogate sample is a pooled sample of negative male urine that is spiked with known concentrations of CT cell lysate. Again, because E2 involves lysis, it does not mimic patient sample processing and is therefore less preferable. Even though we can control the low-, medium-, and high-level spiking for both E1 and E2, the spiked artificial analyte is not preferable, so we are moving to the next level.

Appendix C. (Continued)

Case Study 1

Hierarchy: CT - Urine Specimen

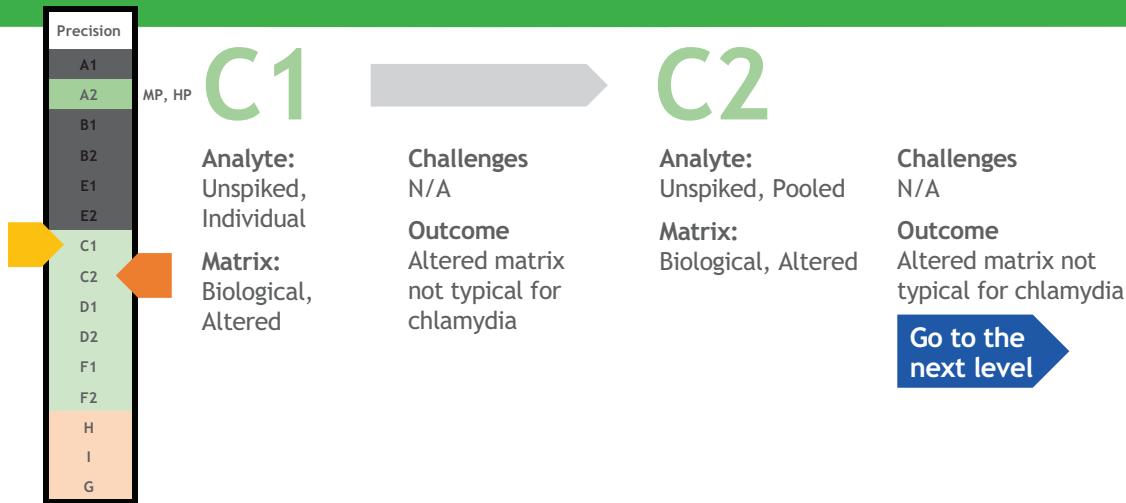


Figure C6. Analyte: Unspiked, Individual; Matrix: Altered, Biological

Now we are at the C samples, where the analyte is unspiked, but now the matrix is altered. The C1 sample is an individual patient sample, and the C2 sample is pooled from more than one patient. The matrix for both C1 and C2 is biological and altered. There are challenges in preparation of the sample with no analyte, because the matrix can be altered. However, while used for other analytes, such procedures are not typical for CT.

Case Study 1

Hierarchy: CT - Urine Specimen

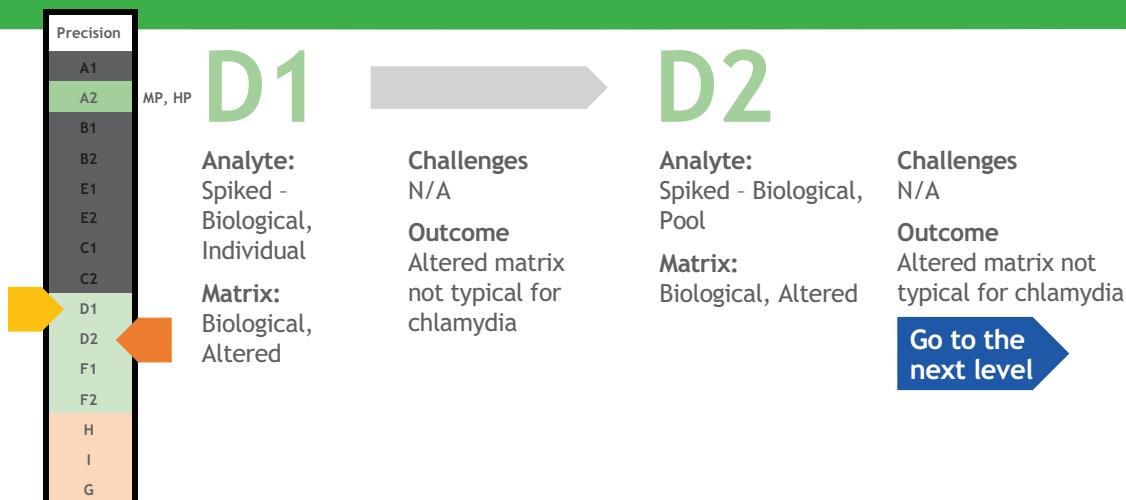


Figure C7. Analyte: Spiked - Biological, Individual; Matrix: Altered, Biological

Appendix C. (Continued)

The matrix is still biological in D samples, but it is altered, and now the analyte has a biological spike. An example of this scenario is stripping the analyte from the matrix prior to spiking. Again, because the altered matrix is not typical for CT, we will go to the next level.

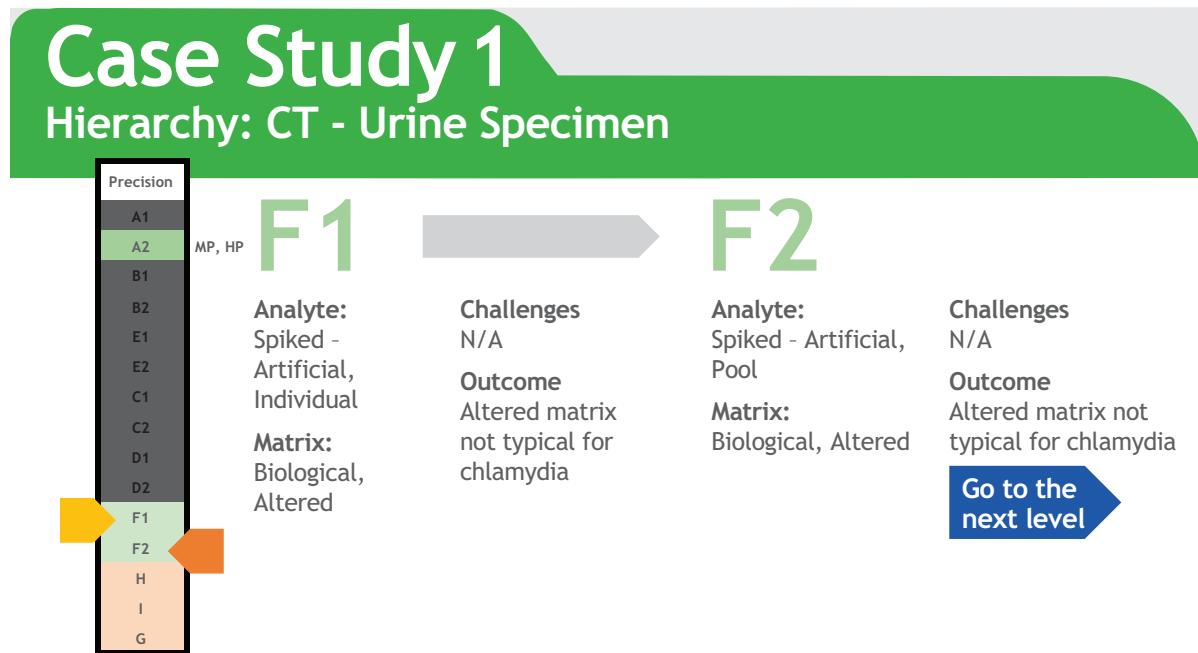


Figure C8. Analyte: Spiked - Artificial, Individual or Pooled; Matrix: Altered, Biological

In the F samples, we still have a biological, but altered, matrix. The analyte is still spiked, but this time it's with an artificial analyte instead of a biological analyte. As we have seen with the C and D samples, an altered matrix is not typical for CT. So we will move past the F samples, and we go to the next level.

Appendix C. (Continued)

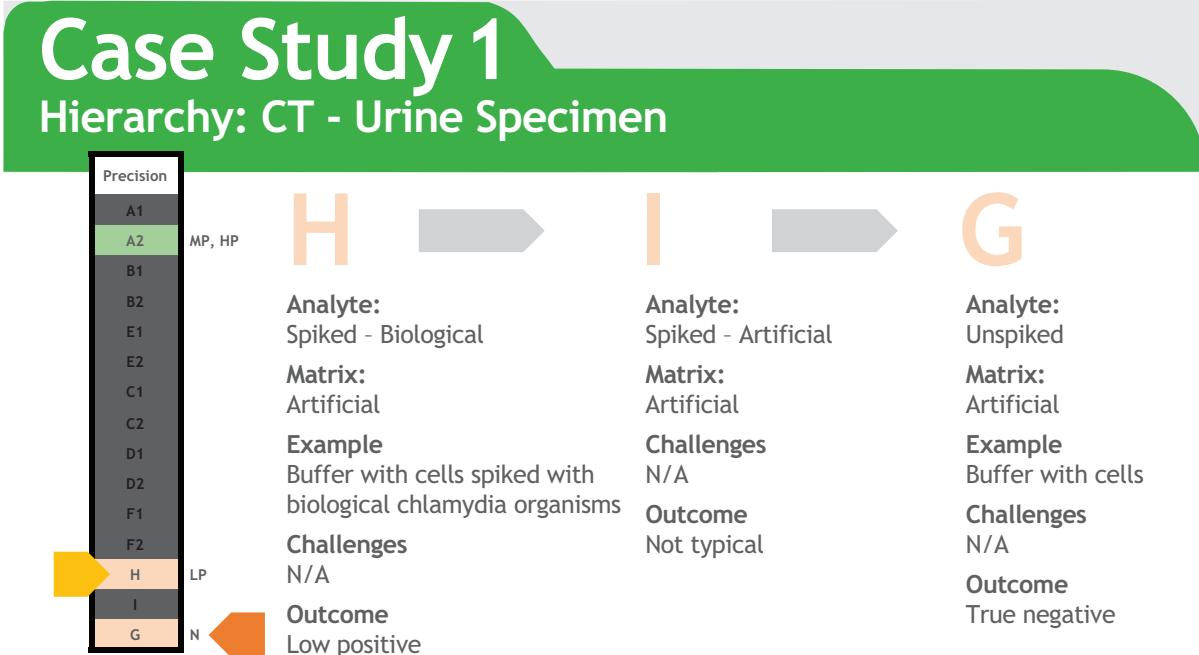


Figure C9. Analyte: Spiked - Biological or Artificial, Individual; Matrix: Artificial

We have arrived at the hierarchy of samples H, I, and G, all of which have an artificial matrix. For H, we spike with a biological analyte, for example, a buffer containing preservative and cells that mimic urine that is then spiked with biological CT organism. Remember that we still need a low-positive sample, which we can create by spiking biological chlamydia organism into a completely negative matrix to create a sample close to the cutoff. This is the composition of sample H and is therefore appropriate for a low positive.

For I, we spike with an artificial analyte. An example of surrogate sample I would be a buffer containing preservative and cells that mimic urine that is then spiked with artificial CT organism. Surrogate sample I is not typical for chlamydia.

G surrogate samples do not have anything spiked into the sample. An example of G would be a buffer containing preservative and cells that mimic urine. Surrogate sample G is appropriate for a true negative. The sample consists of an artificial matrix and is unspiked, so we can be certain that the concentration of chlamydia is zero. A true negative is very important for a precision study.

When we navigate the hierarchy of different types of surrogate samples, we must also prepare a justification of why particular samples were selected. This justification is based on the premise that the best way to select surrogate samples is to select those that mimic the patient samples as closely as possible.

Appendix C. (Continued)

Case Study 1 Conclusion

Chlamydia trachomatis: Precision/Reproducibility Study
Specimen: Urine

Surrogate Sample Type	Description	With Equivalency Testing (An Alternative Surrogate Sample Approach)	Description
Negative (N)	G	Analyte: Unspiked Matrix: Artificial	G
Low positive (LP)	H	Analyte: Spiked - Biological Matrix: Artificial	H
Medium positive (MP)	A2	Analyte: Unspiked, Pooled Matrix: Biological, Unaltered	H
High positive (HP)	A2	Analyte: Unspiked, Pooled Matrix: Biological, Unaltered	H - with regard to matrix

Figure C10. Case Study 1: Urine Specimen - Conclusion

- Negative: we arrived at surrogate sample G, an artificial matrix (buffer and cells that mimic urine) not spiked with any analyte.
- Low positive: we arrived at surrogate sample H, an artificial matrix that is completely negative and spiked with biological analyte.
- Medium positive: we arrived at surrogate sample A2, an unaltered, biological matrix that is pooled and unspiked.
- High positive: we arrived at surrogate sample A2.

Remember that we discussed that we should have pooled patient samples for medium-positive and high-positive samples, but if the volume of the patient samples is small, we can spike sample type G, the artificial negative sample. The artificial negative sample does not have limitations with volume. We can use this sample type G and then spike a biological analyte in a concentration far from the cutoff so that we have 100% of positive results, making this a medium-positive sample. We can spike a very high level of biological analyte to create a high-positive sample. If you would like to use artificial samples (G) even though it is possible to have patient samples that are true negatives (A2), an additional study should be performed to show an equivalency of these matrixes. For the example of CT assay precision, this is not the case, as A2 had a risk for not being a true negative. Therefore, surrogate sample type G was selected for the true-negative sample.

C5 Study-Specific Hierarchy: Vaginal Swab Specimen

We are moving to the second specimen type for a qualitative CT precision study, navigating the hierarchy for a vaginal swab specimen.

Appendix C. (Continued)

We still need the four samples:

- **Negative**, denoted as N
- **Low positive**, denoted as LP, in which the concentration is close to the cutoff
- **Medium positive**, denoted as MP, in which the concentration is nearly 100% positive
- **High positive**, denoted as HP, in which the concentration is nearly at the upper limit of the working range of the assay

Case Study 1

Hierarchy: CT - Vaginal Swab Specimen

Precision

A1
A2
B1
B2
E1
E2
C1
C2
D1
D2
F1
F2
H
I
G

A1

Analyte:
Unspiked, Individual

Matrix:
Biological, Unaltered

Example
Vaginal or cervical swab from one patient

Challenges

- May have insufficient volume to perform a large study
- Longer lead time to collect sufficient samples for the study

Outcome

Go to the next level

Figure C11. Analyte: Unspiked, Individual; Matrix: Unaltered, Biological

Here we have A1, a vaginal or cervical swab sample from an individual. The A1 matrix is biological and unaltered, and the analyte is unspiked. As with urine samples, an individual patient sample here may present volume limitations. If we combine different samples from the same patient in an effort to address the volume limitations, we would need a lot of time to collect all of those samples. Due to the challenges of using an individual patient sample to conduct a precision study, let's go to the next level.

Appendix C. (Continued)

Case Study 1

Hierarchy: CT - Vaginal Swab Specimen

Precision	
A1	
A2	MP, HP
B1	
B2	
E1	
E2	
C1	
C2	
D1	
D2	
F1	
F2	
H	
I	
G	

A2

Analyte: Unspiked, Pooled

Matrix: Biological, Unaltered

Example
Pooled swab samples (swabs and cells extracted individually, then pooled)

Challenges

- Impractical; need to pool a very large number of samples to perform the test
- N - Pool may not be true negative
- LP - With regard to low-positive sample, unable to target concentrations close to the cutoff
- MP & HP - May be appropriate, but depends on available volume
- All - This surrogate sample type may be difficult to evaluate in a multiplex assay

Outcome
Appropriate for MP and HP; however, may need to be spiked; otherwise, go to next level

Go to the next level

Figure C12. Analyte: Unspiked, Pooled; Matrix: Unaltered, Biological

In A2, the sample is unspiked and pooled from more than one patient. The matrix is biological and unaltered. There are several challenges to consider for the samples in the precision study:

- The negative might not be a true negative. We already discussed how it is difficult to check whether this sample really has zero analyte concentration because all tests have a limit of detection, so there is some possibility of a false-negative result. Patients with a CT infection are very frequently asymptomatic, so we really don't have a mechanism to definitely know whether this sample is a true negative.
- For a low-positive sample, pooling can make it difficult to achieve a concentration that is close to the cutoff.
- For medium- and high-positive samples, it may be difficult to pool a large enough volume. However, it is possible.

To summarize, we see that the A2 surrogate may be appropriate for medium and high positives. We have made a note by placing MP and HP to the right of A2 in the hierarchy on the left side of the graphic. Regarding the LP and the negative, A2, does not give us what we need. Therefore, we need to go to the next level, B1.

Appendix C. (Continued)

Case Study 1

Hierarchy: CT - Vaginal Swab Specimen

Precision
A1
A2
B1
B2
E1
E2
C1
C2
D1
D2
F1
F2
H
I
G

MP, HP

B1

Analyte:
Spiked - Biological,
Individual

Matrix:
Biological, Unaltered

Example
Material from one patient
sample with high analyte
concentration spiked into
another patient sample

Challenges

- All - May have insufficient volume to perform a large study
- All - Low throughput
- All - Need to spike an amount that is close to the cutoff. The exact amount of high analyte concentration may be difficult to determine

Outcome

[Go to the next level](#)

Figure C13. Analyte: Spiked - Biological, Individual; Matrix: Unaltered, Biological

The sample is from an individual patient and spiked with biological analyte. The matrix is biological and unaltered. An example of this B1 surrogate is material from a patient sample with a high CT concentration spiked into another patient sample. In terms of challenges, we'd expect volume limitations and difficulty determining the dilution of analyte to spike in to achieve a concentration close to the cutoff, so let's go to the next level.

Case Study 1

Hierarchy: CT - Vaginal Swab Specimen

Precision
A1
A2
B1
B2
E1
E2
C1
C2
D1
D2
F1
F2
H
I
G

MP, HP

LP, MP, HP

B2

Analyte:
Spiked -
Biological,
Pooled

Matrix:
Biological,
Unaltered

Example

Pool of known assay-negative patient samples combined and spiked with a known assay-positive patient sample.

A sample spiked with biological CT organisms reproduced by cell culture is a biological analyte.

Challenges

- N - Pool may not be a true negative
- LP & HP - Sufficient patient sample volume may not be available
- LP - Need to spike an amount that is close to the cutoff. The exact amount of high analyte concentration may be difficult to determine

Outcome

- For HP & MP, if unable to accomplish with A2, B2 is appropriate
- Appropriate for LP

[Go to the next level](#)

Figure C14. Analyte: Spiked - Biological, Pooled; Matrix: Unaltered, Biological

Appendix C. (Continued)

In B2, the sample is pooled from more than one patient and spiked with a biological analyte. The matrix is biological and unaltered. An example of this B2 surrogate is a pool of known negative patient samples spiked with urine from a patient with high CT values or spiked with CT organism reproduced by cell culture. We encounter some challenges for B2 vaginal swabs:

- The negative might not be a true negative.
- For low- and high-positive samples, there may be volume limitations.
- For a low-positive sample, it may be difficult to determine the appropriate analyte dilution or concentration for spiking to achieve a concentration close to the cutoff, because a relationship between signal and concentration in qualitative tests is not as simple as for quantitative tests.

While B2 is a possible option for LP (and we have made a note of this on our hierarchy next to B2), it will likely be very challenging. Therefore, we need to continue on in our hierarchy for the LP. B2 can be appropriate for medium-positive and high-positive samples if you are unable to achieve those concentrations with A2; we have made a note of this on our hierarchy next to B2. We still need to find the right surrogate type for the negative, so let's go to the next level.

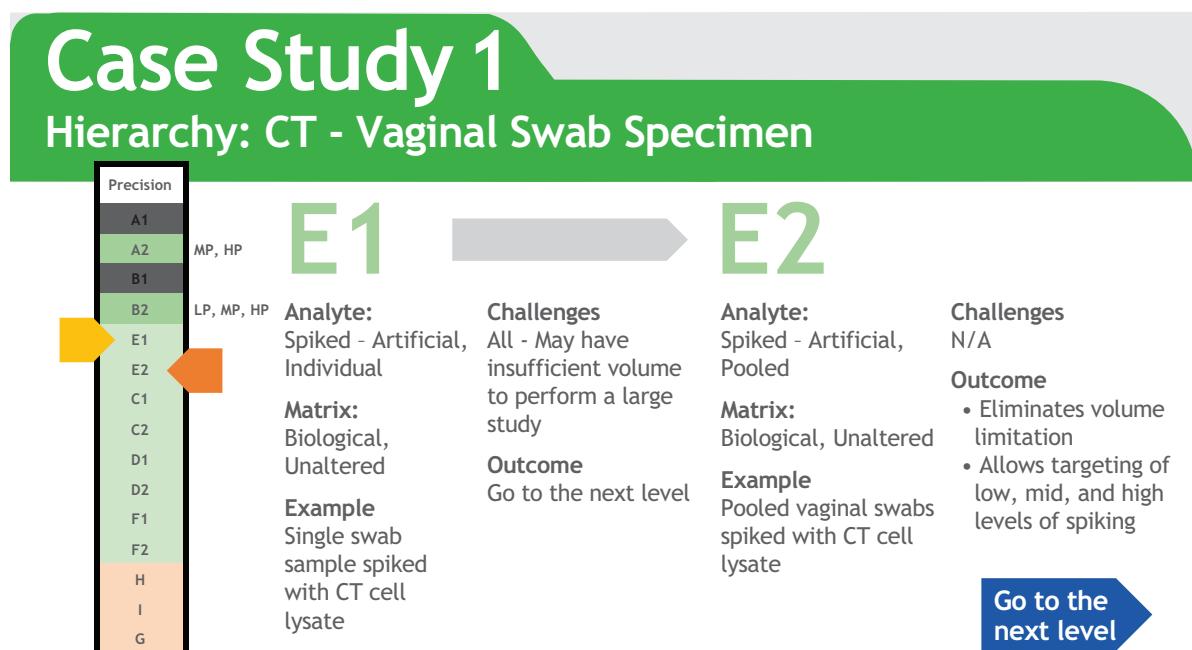


Figure C15. Analyte: Spiked - Artificial, Individual or Pooled; Matrix: Unaltered, Biological

Now we are moving on to the E samples, in which the sample is from an individual patient and spiked with an artificial analyte. The matrix is biological and unaltered. An example of an E1 surrogate sample is a single patient swab specimen spiked with CT cell lysate. The challenge with E1 is that we may have volume limitations that prevent us from performing a large study.

Appendix C. (Continued)

In E2, the sample is pooled from more than one patient and spiked with an artificial analyte. The matrix is biological and unaltered. An example of an E2 surrogate sample is pooled vaginal swabs spiked with CT cell lysate. Using E2 eliminates volume limitations, but it uses artificial analyte (CT cell lysate) and is not preferable because it doesn't entirely mimic all steps for patient sample processing. Also, we have already selected A2 or B2 as our LP, MP, and HP, all of which are higher up on the hierarchy, so we can move on to the next level.

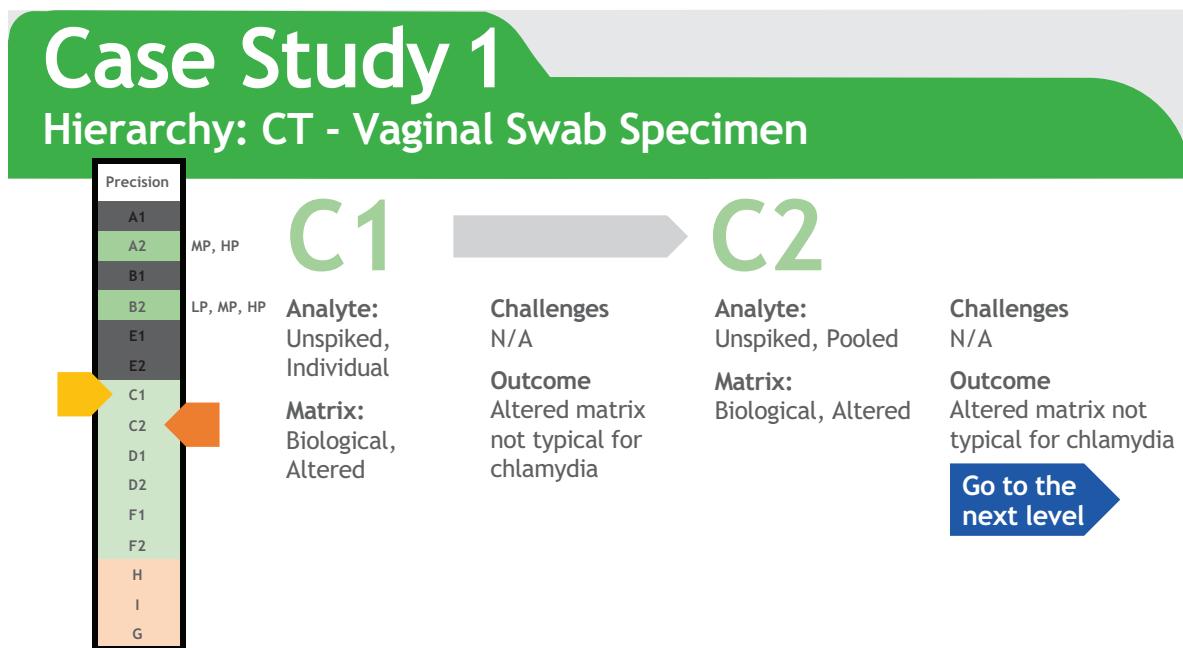


Figure C16. Analyte: Unspiked, Individual; Matrix: Altered, Biological

Now we are at the C samples, in which the analyte is unspiked and now the matrix is altered. C1 is an individual patient sample, and C2 is a pool of samples from more than one patient. The matrix for both C1 and C2 is biological and altered. As we have discussed, there are challenges for preparing the samples without the analyte of interest, so an altered matrix is not typical for CT.

Appendix C. (Continued)

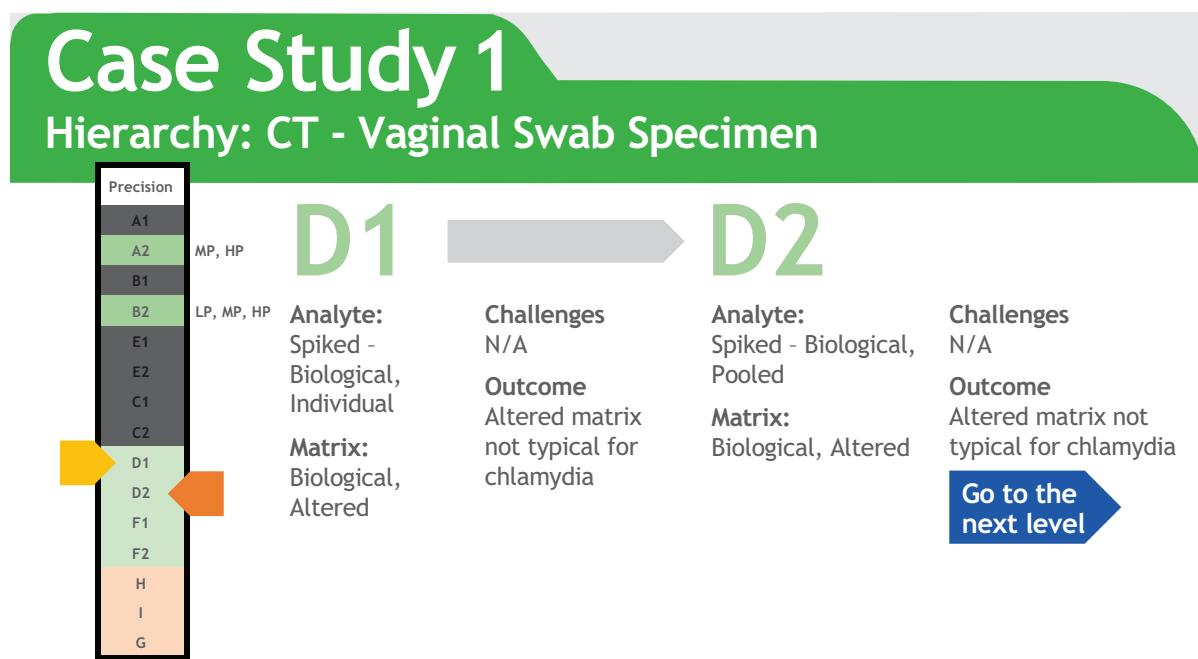


Figure C17. Analyte: Spiked - Biological, Individual; Matrix: Altered, Biological

Now we are at the D samples, in which the matrix is still biological but altered and now the spiked analyte is biological. The matrix for D1 and D2 is an altered biological matrix. Again, because the altered matrix is not typical for CT, we will go to the next level.

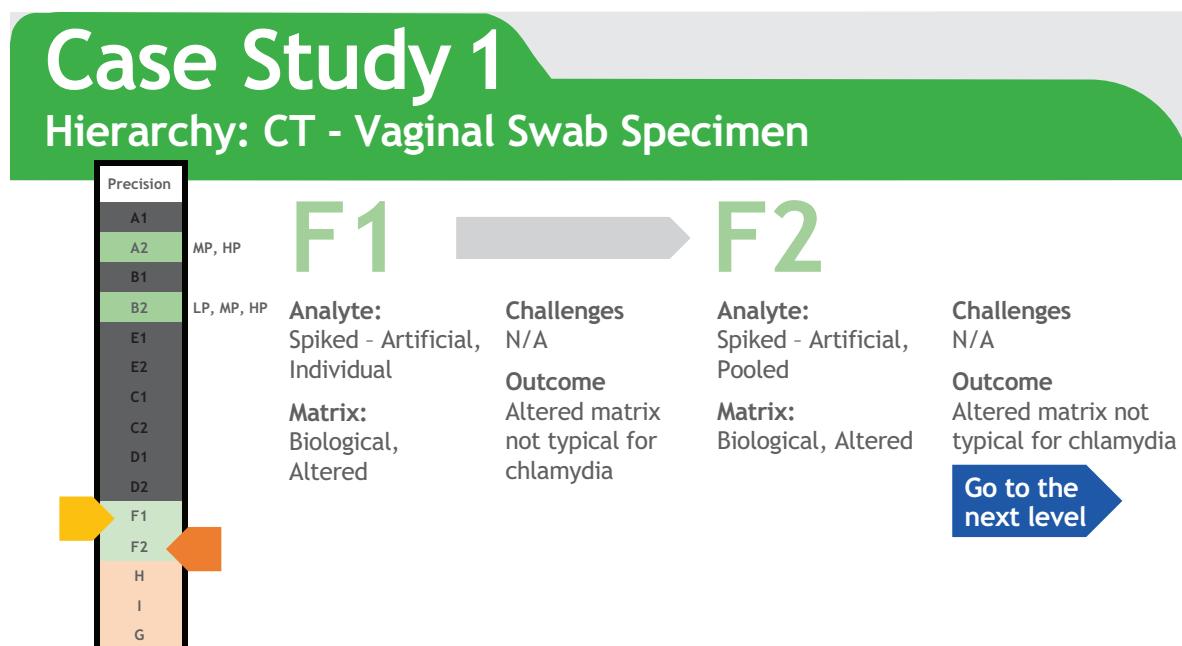


Figure C18. Analyte: Spiked - Artificial, Individual; Matrix: Altered, Biological

Appendix C. (Continued)

In the F samples, we still have a biological but altered matrix. The sample is still spiked, but this time with artificial analyte instead of biological analyte. As we have learned with the C and D samples, an altered matrix is not typical for CT, so we will move past the F samples and to the next level.

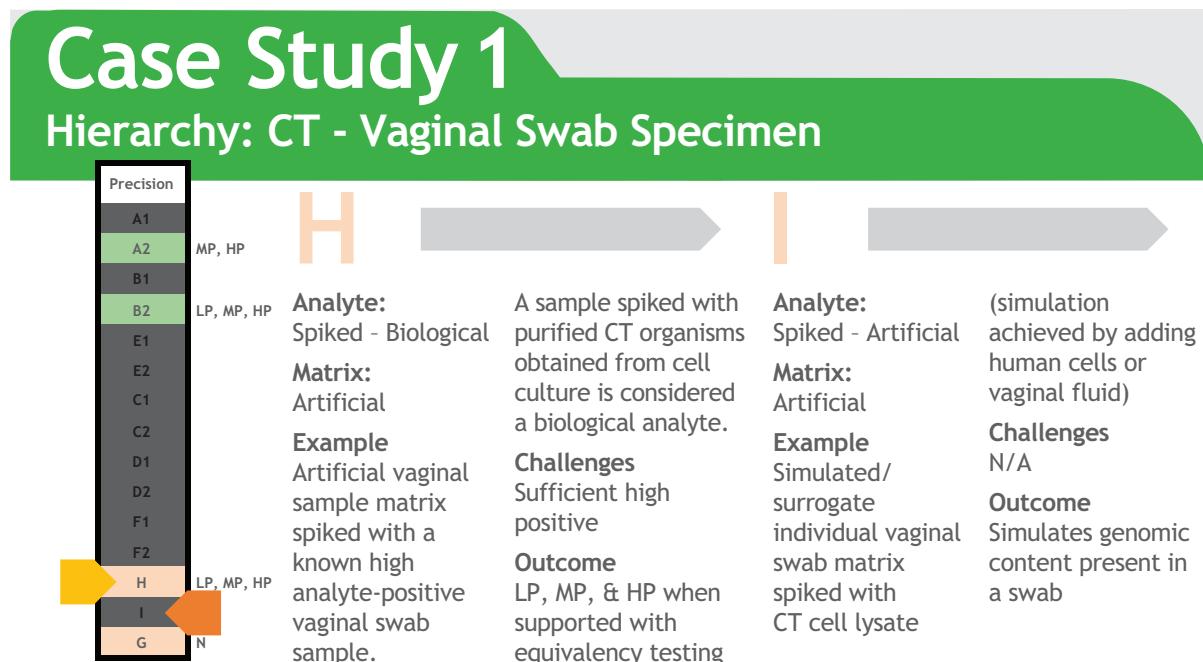


Figure C19. Analyte: Spiked - Biological or Artificial; Matrix: Artificial

We have arrived at samples H, I, and G, all of which have an artificial matrix. The artificial matrix could be buffer with transport media and cells that can be found in vaginal swabs.

For H, we spike with a biological analyte. An example of surrogate sample H would be an artificial vaginal sample matrix spiked with a known high-CT positive vaginal swab sample. Also, we can spike with purified CT obtained from cell culture, because we consider it a biological analyte. Surrogate sample H may be appropriate for the LP. A2 surrogate samples are the best for MP, and either A2 or B2 are the best for HP. However, if there are data to support equivalency testing of B2 and H, then H samples can be also used for MP and HP. We have noted this on the hierarchy next to H.

For I, we spike with an artificial analyte. An example of surrogate sample I would be a simulated surrogate individual vaginal swab matrix spiked with CT cell lysate. We discussed previously that spiking with CT cell lysate does not mimic all processing steps of the patient samples. Sample type I, while it simulates genomic content, is not acceptable because cell culture can be used for CT testing.

Appendix C. (Continued)

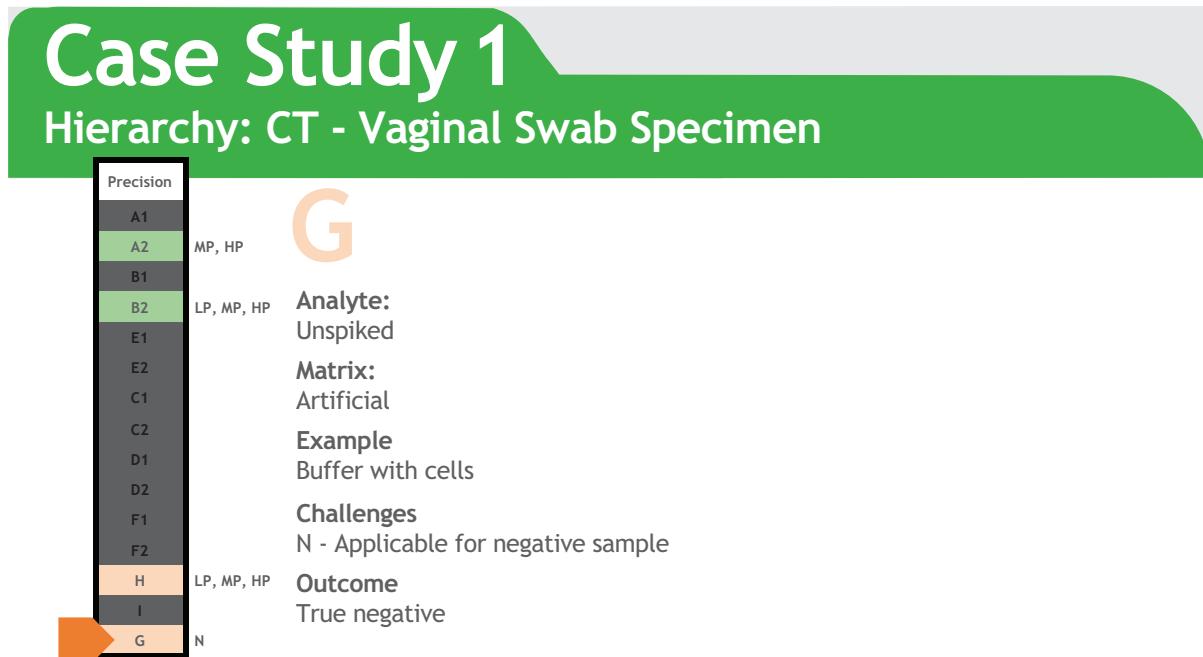


Figure C20. Analyte: Unspiked; Matrix: Artificial

For G, we do not spike in any analyte; it is a negative matrix. An example of surrogate sample G would be a buffer containing cells that mimic those found in a vaginal swab. Surrogate sample G is appropriate for the true negative and we have made a note on the hierarchy next to G.

Appendix C. (Continued)

C6 Conclusion

Case Study 1 Conclusion

Chlamydia trachomatis: Precision/Reproducibility Study
Specimen: Vaginal Swab

Surrogate Sample Type	Description	With Equivalency Testing (An Alternative Surrogate Sample Approach)	Description
Negative (N)	G	G Analyte: Unspiked Matrix: Artificial	Analyte: Unspiked Matrix: Artificial
Low positive (LP)	B2, H	B2 Analyte: Spiked - Biological, Pooled Matrix: Biological, Unaltered H Analyte: Spiked - Biological Matrix: Artificial	H Analyte: Spiked - Biological Matrix: Artificial
Medium positive (MP)	A2, B2	A2 Analyte: Unspiked, Pooled Matrix: Biological, Unaltered B2 Analyte: Spiked - Biological, Pooled Matrix: Biological, Unaltered	H Analyte: Spiked - Biological Matrix: Artificial
High positive (HP)	A2, B2	A2 Analyte: Unspiked, Pooled Matrix: Biological, Unaltered B2 Analyte: Spiked - Biological, Pooled Matrix: Biological, Unaltered	H - with regard to matrix Analyte: Spiked - Biological Matrix: Artificial

Figure C21. Case Study 1 - Vaginal Swab Conclusion

For this case, a CT precision study using vaginal swab specimens, we needed four samples:

- Negative sample: we arrived at surrogate sample G, an artificial matrix not spiked with any analyte.
- Low positive: there are two options; the preferred option is B2, pooled samples in an unaltered biological matrix spiked with biological analyte (CT organism obtained from cell culture). If B2 proves challenging, we can use H, which is a negative artificial matrix spiked with biological CT.
- Medium positive and high positive: again, we have two options; it is best to have pooled patient samples (A2), but if that proves challenging, we can use B2.

As we discussed, this approach can present volume limitations (especially for the high-positive sample). Therefore, surrogate sample H is an alternative (with equivalency testing), in which we have an artificial matrix and spike in biological analyte to prepare the low-positive, medium-positive, and high-positive samples.

Appendix D. Creatinine: Precision/Reproducibility

Appendix D contains excerpts from Medical Device Innovation Consortium (MDIC). *Surrogate Sample Framework Harmonized Education Case Study: Creatinine Precision/Reproducibility*. Arlington, VA: Medical Device Innovation Consortium; 2019, with permission, and has been maintained verbatim from the source material.

NOTE: This case study is written in a conversational tone.

D1 Introduction

This is the second of five case studies intended to demonstrate how to apply the Surrogate Sample Framework. This is another precision or reproducibility study with fresh urine specimens, but this time the analyte is creatinine. Unlike the previous case study for a qualitative test, this one is a quantitative test.

D2 Principles

There are seven principles applicable to this case study.

D2.1 Principle 1 - Identify the Objective for Using Surrogate Samples for This Study

The precision study design for a quantitative test is relatively similar to that of a qualitative test; the concentrations are different, but the study design is the same. For a reproducibility study, we need to have, for example, three sites. We need to test the same sample for a relatively long period of time, for example, for five days, or maybe even for 20 days. We also need to send this sample to different sites, so the samples should be stable during the reproducibility study and also through shipping and storage. Because we need to have a lot of replicates, we need to have enough volume. These volume and stability requirements are difficult to meet with fresh urine.

D2.2 Principle 2 - Identify Critical Factors Related to the Performance Study Type

In the precision study for qualitative tests, we need to have a negative sample and samples close to the cutoff. However, quantitative tests have measuring intervals, so we need samples to assess precision at the lower and upper limits of the measuring interval, as well as close to the medical decision level (MDL). If we have two medical decision levels for the test, we need to have two unique concentrations: one that is close to one medical decision level and a second concentration that is close to the second medical decision level.

Appendix D. (Continued)

D2.3 Principle 3 - Consider Factors Based on Whether the Test Is Qualitative, Quantitative, or Semi-Quantitative

This is a quantitative test, so we need to think about what that means for our surrogate sample selection. We will need to have concentrations that should include at least one concentration close to each medical decision level and concentrations close to the lower and upper limits of the measuring interval.

D2.4 Principle 4 - Consider the Effect of Sample Processing Such as Extraction Method, If Applicable

For this assay and this type of specimen (fresh urine), urine surrogate samples have been considered representative samples of the patient samples.

D2.5 Principle 5 - Evaluate Property Characteristics of the Matrix

A pool of a single individual's samples does not have any problems with regard to matrix. If we pool urine from different patients, this sample will present as an "average" matrix. If there are not any artificial effects, this "average" matrix can be considered a representative urine sample. In addition, pooled samples from multiple patients may be a better representation of the typical patient sample than a pool of samples from one patient, because there is a chance that a single patient can be atypical.

D2.6 Principle 6 - Evaluate Property Characteristics of the Analyte

Unlike the previous study for chlamydia, in which we had to consider different types of CT, creatinine is a well characterized analyte with no biological variants of interest.

D2.7 Principle 7 - Assess Analyte and Matrix Combinations of Surrogate Samples

The urine matrix remains unchanged except for the potential addition of a small volume of high-concentration creatinine stock solution. The creatinine in the stock is chemically identical to that found in patient urine specimens.

D3 Navigating the Hierarchy

We will be navigating the hierarchy for a precision/reproducibility study where the analyte is creatinine. It is a quantitative test, and the specimen is fresh urine. The surrogate samples we need for the study are various concentrations of positive samples. By "positive", we mean samples that contain analyte. These should include concentrations close to the lower and upper limits of the measuring interval and one concentration near each medical decision level.

Appendix D. (Continued)

Case Study 2

Hierarchy: Creatinine - Urine Specimen

Precision
A1
A2
B1
B2
E1
E2
C1
C2
D1
D2
F1
F2
H
I
G

A1

Analyte: Unspiked, Individual

Matrix: Biological, Unaltered

Example
Urine sample from one patient, male or female

Challenges
Need for large volume with ranges of concentrations. A single donor might have a characteristic that will compromise interpretation.

Outcome

Go to the next level 

Figure D1. Analyte: Unspiked, Individual; Matrix: Unaltered, Biological

A1 is the first sample type to be evaluated. The sample is unspiked and from an individual patient. The matrix is biological and unaltered. An example of an A1 sample is a urine sample from one patient, male or female. With A1, we will need to collect large volumes with different concentrations, which could be a challenge. Also, since we are dealing with only one patient, there is some chance that the sample might have atypical characteristics, such that estimation of imprecision in the surrogate will not represent the imprecision for typical samples.

Appendix D. (Continued)

Case Study 2

Hierarchy: Creatinine - Urine Specimen



A2

Analyte: Unspiked, Pooled

Matrix: Biological, Unaltered

Example
Pooled urine from male and/or female samples

Challenges
Need for large volumes with ranges of concentrations

Outcome

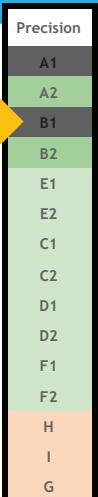
Go to the next level 

Figure D2. Analyte: Unspiked, Pooled; Matrix: Unaltered, Biological

In A2, the sample is unspiked and pooled. The matrix is biological and unaltered. An example of an A2 surrogate is pooled urine from male or female samples. With A2, we will again need to collect large volumes with the necessary ranges of concentrations, which could be a challenge.

Case Study 2

Hierarchy: Creatinine - Urine Specimen



B1

Analyte: Spiked - Biological, Individual

Matrix: Biological, Unaltered

Example
Patient sample spiked with creatinine

Challenges
A single donor might have a characteristic that will compromise interpretation.

Outcome

Go to the next level 

B2

Analyte: Spiked - Biological, Pooled

Matrix: Biological, Unaltered

Example
Pooled urine spiked with creatinine

Challenges
Need for large volumes with ranges of concentrations

Outcome
P - Achieved large volume and various concentrations

STOP here 

Figure D3. Analyte: Spiked - Biological, Individual or Pooled; Matrix: Unaltered, Biological

Appendix D. (Continued)

In B1, the sample is from an individual patient and spiked with biological analyte. The matrix is biological and unaltered. An example of this B1 surrogate is a patient sample spiked with creatinine. As was the case with the previous individual samples, a single patient sample does not have enough volume for multiple replicates.

For B2, the sample is pooled from more than one patient and the sample is spiked with a biological material. The matrix is biological and unaltered. An example of this B2 surrogate is a pooled urine sample spiked with creatinine. In a precision study, we need large volumes with particular levels of concentrations. If we pool patient samples and see that the concentration is close to what we need, then we do not need to do any additional steps. If the concentration is too low, we will spike this sample; this is surrogate B2. By spiking the creatinine, we can have good control over the concentration.

D4 Conclusion

Case Study 2 Conclusion		Creatinine: Precision/Reproducibility Study Specimen: Urine
Surrogate Sample Type	Description	
Positive (P) at various concentrations	B2	Analyte: Spiked - Biological, Pooled Matrix: Biological, Unaltered

Figure D4. Case Study 2 - Conclusion

We will stop here and select B2. We have noted the selection in the hierarchy in Figure 4.

For this quantitative creatinine assay with urine specimens, we wanted to select positive surrogate samples at various concentrations: close to the lower limit of the measuring interval (low concentration), close to the medical decision levels, and close to the upper limit of the measuring interval (high). We have selected surrogate sample B2 because it is possible to have a matrix that is biological and without alteration. We plan to use pooled patient samples, so we may need to spike analyte to reach desired sample concentrations for some concentration levels, for example, close to the MDL. Regarding surrogate sample selection, you can see that there is a big difference between the two precision studies: a qualitative assay (the chlamydia case study) and a quantitative assay (this creatinine case study), even though both studies are precision studies.

Appendix E. Total Prostate Specific Antigen: Precision/Reproducibility

Appendix E contains excerpts from Medical Device Innovation Consortium (MDIC). *Surrogate Sample Framework Harmonized Education Case Study: Total Prostate Specific Antigen - Precision/Reproducibility*. Arlington, VA: Medical Device Innovation Consortium; 2019, with permission, and has been maintained verbatim from the source material.

NOTE: This case study is written in a conversational tone.

E1 Abbreviations and Acronyms

MDL medical decision level
PSA prostate specific antigen

E2 Introduction

This is the fourth of five case studies intended to demonstrate how to apply the Surrogate Sample Framework. This is another precision or reproducibility study, but this time the analyte is total prostate specific antigen (total PSA), and the specimen type is serum. This case study is interesting because you will see how different types of tests can affect surrogate sample preparation. Total PSA is a quantitative test.

E3 Principles

There are seven principles applicable to this case study.

E3.1 Principle 1 - Identify the Objective for Using Surrogate Samples for This Study

Surrogates need to be used, as a sample from a single individual does not provide enough volume to create the number of replicates needed for the precision study. It should be stable over the time needed for the precision study, which could be five or 20 days.

E3.2 Principle 2 - Identify Critical Factors Related to the Performance Study Type

For precision studies, we know that having enough volume and having a stable sample are key factors. Stability encompasses not only analytical stability, the time to perform the study, but also shipping stability for studies that are conducted at multiple sites.

Appendix E. (Continued)

E3.3 Principle 3 - Consider Factors Based on Whether the Test Is Qualitative, Quantitative, or Semi-Quantitative

Quantitative tests have measuring intervals. For example, we know that the total PSA test can measure from 0.5 to 100 ng/mL. Precision needs to be understood throughout the entire measuring interval. Therefore, a concentration close to the lower limit of the measuring interval and another concentration close to the upper limit of the measuring interval are needed.

In addition, the medical decision levels (MDLs) for total PSA are 2.5 ng/mL, 4 ng/mL, and 10 ng/mL, so samples around those concentrations are needed. The test can measure up to 100 ng/mL, leaving a large gap between 10 ng/mL and 100 ng/mL. Therefore, a sample in between 10 ng/mL and 100 ng/mL is necessary, so we might have a sample at 40 ng/mL.

E3.4 Principle 4 - Consider the Effect of Sample Processing Such as Extraction Method, If Applicable

Remember that for precision, there is a concern about stability over the test interval and during shipping/storage. Preservatives will likely be added in order to maintain stability; hence, it is necessary to be aware of the effect of those preservatives on the sample.

E3.5 Principle 5 - Evaluate Property Characteristics of the Matrix

A pool of serum can provide a relatively uniform matrix for assessing reproducibility. If serum is pooled from different patients, large changes in matrix are not anticipated. To achieve the concentrations needed for this precision study, especially near the upper limit of the measuring interval, we can supplement with PSA and use the World Health Organization standard.

E3.6 Principle 6 - Evaluate Property Characteristics of the Analyte

Total PSA represents a combination of free PSA and complexed PSA. World Health Organization (WHO) material, this material has a 90-to-10 ratio of complexed PSA to free PSA, if added. If the patient sample is supplemented with WHO material, this ratio can be changed. If samples from the different patients are pooled, the ratio of complexed PSA to free PSA can also be changed. In the precision study, the PSA ratio in the sample is not as important as having a sample that is close to the MDLs. It is okay if the samples close to the different MDLs have different ratios, because what is really being evaluated is the precision of the total PSA.

E3.7 Principle 7 - Assess Analyte and Matrix Combinations of Surrogate Samples

It is probably very difficult to find a patient sample with very high values of total PSA, so the World Health Organization (WHO) material will be used. When spiked with WHO standard, as little as possible should be used to most closely mimic a “native” patient sample. Due to this goal, a patient sample with a natural analyte concentration as high as possible is needed, and then spike in the smallest amount of World Health Organization standard as possible. However, as mentioned previously, if supplemented with PSA, do expect that ratio of free to complexed PSA to change, but the ratio is not important for the precision study of the total PSA.

Appendix E. (Continued)

E4 Navigating the Hierarchy

The hierarchy for a precision/reproducibility study where the analyte is total PSA is navigated. It is a quantitative test, and the specimen is serum. The samples needed for the study are various concentrations of positive samples. “Positive” means samples that contain analyte. Samples should include concentrations close to the lower and upper limits of the measuring interval and concentrations near each medical decision level.

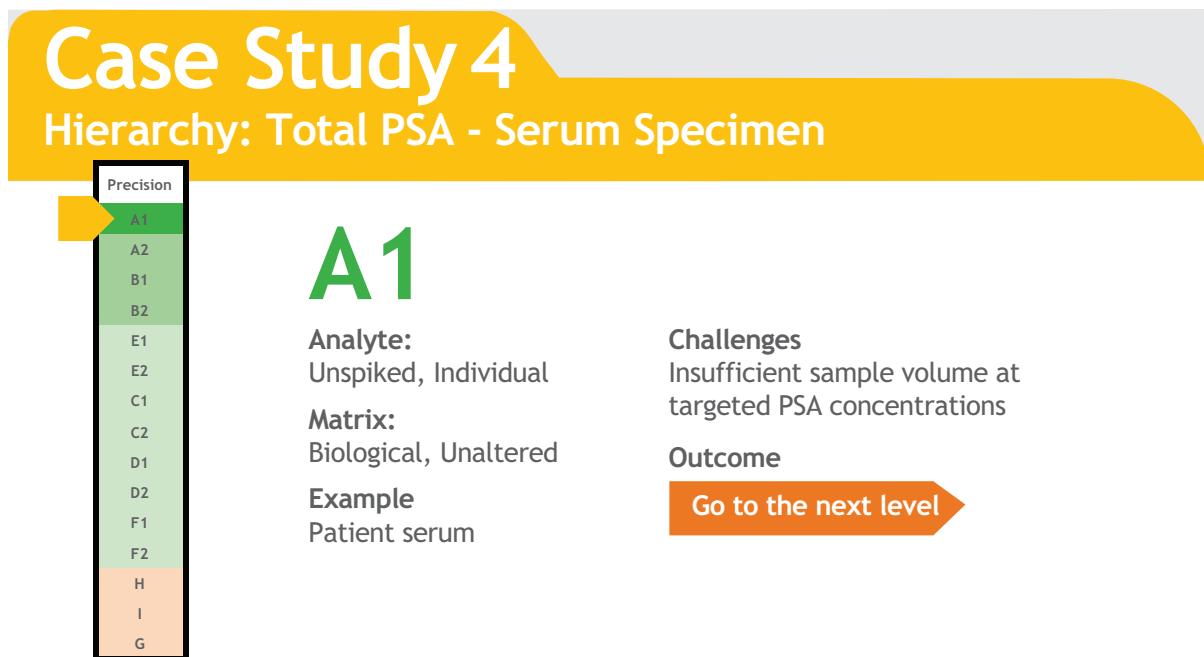


Figure E1. Analyte: Unspiked, Individual; Matrix: Unaltered, Biological

A1 is first. The sample is unspiked serum from an individual patient. The challenge with A1 is that we don't have enough sample volume for this type of study.

Appendix E. (Continued)

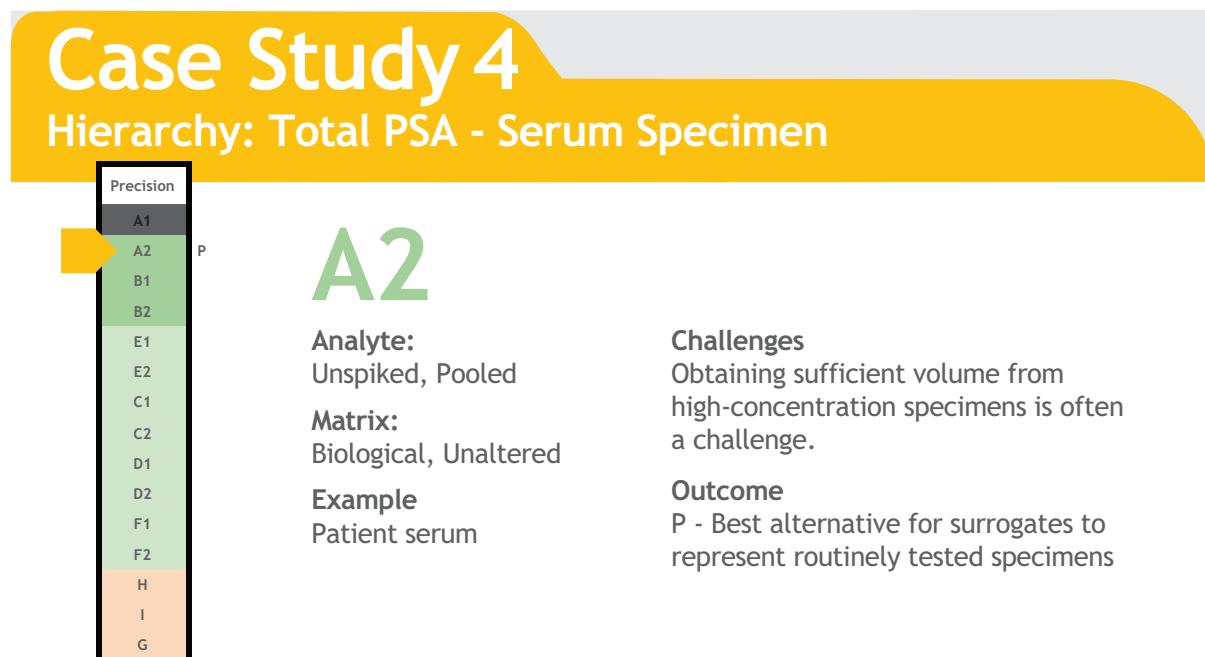


Figure E2. Analyte: Unspiked, Pooled; Matrix: Unaltered, Biological

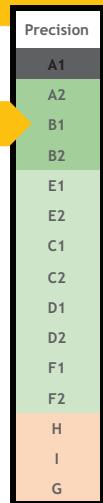
In A2, the sample is a pool of serum samples from different patients. Use pooled patient samples to obtain the volume necessary for a reasonable number of replicates for a reproducibility study for five days (within-laboratory precision study) or 20 days (reproducibility precision study) and do not do any spiking. With regard to samples close to the MDLs of 2.5, 4.0, and 10.0 ng/mL, we see that we can find patient samples with values close to MDLs and then pool the patient samples with values close to the corresponding MDL (eg, pool patient samples with values close to 4.0 ng/mL to create a mini-pool with a value close to 4.0 ng/mL). If we can't find patient samples with values close to the MDLs, we will need to pool patient samples with different total PSA values. Note that patient samples with high values can be pooled with another patient sample with a much lower value to obtain concentrations close to the MDL (eg, a patient sample with a value of 1 ng/mL and a patient sample with a value of 20 ng/mL can be pooled in the same proportion to have a mini-pool with a value close to 10 ng/mL). The values of the patient samples will determine the proportion in which the samples are pooled.

Please note that if there is a need to have a sample with a total PSA value close to the lower limit of the measuring interval, for example, 1 ng/mL, it may be difficult to find these samples in the intended use population of males 50 years or older. For the precision study, use samples from the population of males who are younger, because they usually have a very low value for total PSA. Pool these patient samples to create an A2 sample that is close to the lower limit of the measuring interval. Therefore, A2 is appropriate for the positive surrogate sample, as noted in Figure E2.

Appendix E. (Continued)

Case Study 4

Hierarchy: Total PSA - Serum Specimen

**B1**

Analyte: Spiked - Biological, Individual
Matrix: Biological, Unaltered
Example Patient serum

Challenges
 Insufficient sample volume at targeted PSA concentrations
Outcome

Go to the next level

B2

Analyte: Spiked - Biological, Pooled
Matrix: Biological, Unaltered
Example Patient serum

Challenges
 Ratio of free to complexed PSA is thought to change.

Outcome
 HP - Supplementing a serum pool with a high-concentration material, such as WHO standard, provides sufficient volume and stability to achieve goals of reproducibility studies.

STOP here

Figure E3. Analyte: Spiked - Biological, Individual or Pooled; Matrix: Unaltered, Biological

In B1, the sample is from an individual patient and spiked with biological analyte. The matrix is biological and unaltered. It is discussed that this sample type will not work for the precision study, because it is really necessary to have a large volume for several replicates.

For B2, the sample is pooled from more than one patient sample, and it is spiked with biological analyte. The matrix is biological and unaltered. In order to spike with biological material, find patient samples with very high values. However, this is not easy. We are stuck in a cycle where we know that to prepare a sample with a very, very high (100 ng/mL) value, patient samples with a very high value need to be found, which is challenging. To overcome this issue, spike with WHO standard material, which can be considered a biological analyte. It is important to remember to spike with as little of the WHO standard as possible to maintain the biological matrix and achieve the desired concentrations, pool patient samples with relatively high concentrations, for example, 20 ng/mL or 30 ng/mL, and then spike the WHO standard.

A challenge with B2 is that PSA supplementation may change the ratio of free to complexed PSA. However, this is not an issue for a precision study. Therefore, surrogate sample B2 is appropriate for the high-positive sample, a sample close to the upper limit of the measuring interval. We will stop here and select B2. The selection is noted in Figure E3.

Appendix E. (Continued)

E5 Conclusion

Case Study 4 Conclusion

Total PSA: Precision/Reproducibility Study
Specimen: Serum

Surrogate Sample Type	Description
Positive (P) at various concentrations	A2 Analyte: Unspiked, Pooled Matrix: Biological, Unaltered
High positive (HP)	B2 Analyte: Spiked - Biological, Pooled Matrix: Biological, Unaltered

Figure E4. Case Study 4 - Conclusion

For the precision study of the quantitative total PSA assay using serum specimens, positive samples at various concentrations, including a sample close to the upper limit of the measuring interval, are needed. Samples close to the lower limit of the measuring interval and close to 2.5 ng/mL, 4 ng/mL, 10 ng/mL, and 40 ng/mL are needed. For these samples, surrogate sample A2 is chosen, which is pooled and without spiking. For the sample close to the upper limit of the measuring interval, use surrogate sample B2. The matrix is biological and unaltered, but the surrogate sample is comprised of pooled samples from different patients and spiked with the WHO standard.

Appendix F. Albumin: Method Comparison

Appendix F contains excerpts from Medical Device Innovation Consortium (MDIC). *Surrogate Sample Framework Harmonized Education Case Study: Albumin Method Comparison*. Arlington, VA: Medical Device Innovation Consortium; 2019, with permission, and has been maintained verbatim from the source material.

NOTE: This case study is written in a conversational tone.

F1 Introduction

This is the last of the five case studies intended to demonstrate how to apply the Surrogate Sample Framework. The analyte for this case study is albumin, and the specimen is fresh urine. This time, the analytical study is a method comparison study. In a method comparison study, measure the patient sample with both the candidate test and with a comparator test. The comparator could be a predicate device or reference method, or some other traceable method. Sufficient sample volume for measurement by the candidate and by the comparator method is required.

F2 Principles

There are seven principles associated with this study.

F2.1 Principle 1 - Identify the Objective for Using Surrogate Samples for This Study

Consider the use of samples from different patients with small volume as leftover samples. One replicate for the candidate test and one replicate for the comparator are needed. Depending on the type of specimen, pooling some patient samples in order to have sufficient volume might be necessary.

To preserve biological variability, create a mini-pool of as few patient samples as possible. If it is possible to pool only two patient samples, that is great. If we need to pool three patients, then we need to pool three samples in order to have one surrogate sample in the method comparison study. Note that it is not a good idea to pool, for example, five patients in such a way: mini-pool #1 is samples from patient 1, 2, and 3, and mini-pool #2 is samples from patient 3, 4, and 5. Remember that the need to have mini-pools that mimic different patients (patient 3 should not be used for both pools) is necessary.

Appendix F. (Continued)

For quantitative tests, specimens included in method comparison studies should cover the analytical measuring interval without large gaps in any part of the analytical measuring interval. Surrogate samples may be used to fill gaps in any part of the range in method comparison studies. The surrogate samples used for this purpose should have been previously shown to have similar performance (precision, slope and intercept) to clinical specimens and their concentration should overlap that of native clinical specimens, as described in CLSI document EP35.¹

Another issue in the method comparison study for a quantitative assay is that we need to do regression analysis; we need to ensure we cover the measuring interval. Sometimes, surrogate samples can be placed in the assay range where it is difficult to find patient samples. Try to understand how your surrogate samples can fix a problem with distribution of the samples through the measuring interval.

F2.2 Principle 2 - Identify Critical Factors Related to the Performance Study Type

We already discussed that there can be critical factors such as sample volume, but because we use urine for this particular type of assay, volume will not be a critical factor. We need to have appropriate distribution of the micro-albumin concentration (meaning that we need to cover the measuring interval) so that we can thoroughly evaluate the difference between the two methods over the entire range of concentrations.

F2.3 Principle 3 - Consider Factors Based on Whether the Test Is Qualitative, Quantitative, or Semi-Quantitative

There are no special factors for method comparison studies. Nevertheless, we need to understand that the regression line will be used for evaluating bias at the MDLs, so it is good to have a few samples around MDLs so that the regression line accurately reflects the performance of the two tests around that MDL.

F2.4 Principle 4 - Consider the Effect of Sample Processing Such as Extraction Method, If Applicable

If we need to transport the patient sample for some reason, it is likely that we will have to add preservative to the urine specimen. Therefore, we will also need to add preservative to our surrogate samples. If we add preservatives, it needs to be investigated how the preservative might affect the results of the comparison study. Another factor is the time of the urine collection, which can affect the concentration and stability of the analyte. Therefore, this physiologic factor should be considered in surrogate sample selection and preparation.

F2.5 Principle 5 - Evaluate Property Characteristics of the Matrix

The range of micro-albumin concentrations in urine specimens is easy to obtain for the purpose of method comparison, so there are not any special characteristics of the matrix to consider. As discussed earlier, pools should be created from as few patient samples as possible, to essentially create mini-pools. In this case, there should not be any matrix issues.

Appendix F. (Continued)

F2.6 Principle 6 - Evaluate Property Characteristics of the Analyte

We do not foresee any issues in obtaining the necessary analyte concentrations for this method comparison study.

F2.7 Principle 7 - Assess Analyte and Matrix Combinations of Surrogate Samples

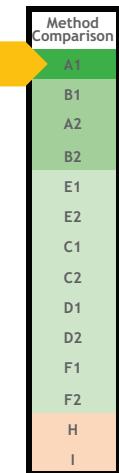
Since the range of micro-albumin concentration in urine is easy to obtain, we would not plan to do any spiking and thus are not concerned about the analyte-matrix combinations of surrogate samples. It is important to assess matrix variability, the scatter of results around the line of best fit. Therefore, individual patient samples should be used as much as possible. If the individual patient sample is limited by volume, it is acceptable to create a mini-pool. Again, the mini-pool should have as few patients as possible, because each surrogate sample should mimic patient samples as much as possible.

F3 Navigating the Hierarchy

We are navigating the hierarchy for a method comparison study where the analyte is albumin. The specimen is urine, and the assay is quantitative.

Case Study 5

Hierarchy: Albumin - Urine sample



A1

Analyte:
Unspiked, Individual
Matrix:
Biological, Unaltered
Example
Patient specimen:
Unspiked, Unaltered

Challenges
No, since the range of micro-albumin concentrations in urine specimens is easily obtained for the purposes of method comparison studies

Outcome
Acceptable



Figure F1. Analyte: Unspiked, Individual; Matrix: Unaltered, Biological

A1 is a sample from an individual patient. The matrix is biological and unaltered. The A1 sample is a patient specimen that is unspiked and unaltered. For an individual patient sample without any spiking, it may be difficult to find samples from a particular concentration range. However, it is usually not difficult to find the desired concentration range with an individual patient specimen for micro-albumin.

If volume is an issue for A1, then mini-pools of as few as possible patient samples can be created to overcome the issue. For this case study, the specimen type is urine, so volume is

Appendix F. (Continued)

not an issue. As such, we need to start with individual patient samples, and each individual patient sample is tested by the candidate assay with one replicate and by the comparator with one replicate.

We will use sample A1, individual patient samples. No surrogate sample is needed. We have now finished navigating the hierarchy for this study.

F4 Conclusion

Case Study 5 Conclusion		Albumin: Method Comparison Study Specimen: Urine
Surrogate Sample Type	Description	
A1	Analyte: Unspiked, Individual Matrix: Biological, Unaltered	

The range of micro-albumin concentration in urine specimens is easily obtained for the purposes of method comparison studies. Therefore, surrogate samples are not generally needed.

Figure F2. Case Study 5 - Conclusion

For this case, an albumin method comparison study using fresh urine specimens, surrogate samples in the method comparison study are typically not needed.

Individual patient samples should be used in the method comparison study because:

1. It is possible to find samples that cover the range of the assay.
2. It is possible to find urine samples of the needed volume, and this volume only needs to be enough for two replicates (one replicate for the candidate assay and one replicate for the comparator).

If there is an issue with sample volume, we discussed the use of mini-pools. If there is an issue with covering the range, we discussed that we may need to spike some amount of the analyte in order to have concentrations to cover the entire range.

In other words, surrogate samples are not generally needed for this assay and this type of specimen (with the rare exception for mini-pools).

Appendix F. (Continued)

Reference for Appendix F

- ¹ CLSI. *Assessment of Equivalence or Suitability of Specimen Types for Medical Laboratory Measurement Procedures*. 1st ed. CLSI guideline EP35. Clinical and Laboratory Standards Institute; 2019.

Appendix G. *Chlamydia trachomatis*: Instrument Carryover

Appendix G contains excerpts from Medical Device Innovation Consortium (MDIC). *Surrogate Sample Framework Harmonized Education Case Study: Chlamydia trachomatis - Instrument Carryover*. Arlington, VA: Medical Device Innovation Consortium; 2019, with permission, and has been maintained verbatim from the source material.

NOTE: This case study is written in a conversational tone.

G1 Introduction

This is the third of five case studies intended to demonstrate how to apply the Surrogate Sample Framework. The analyte is *Chlamydia trachomatis* (CT), and the specimen is a vaginal or endocervical swab. This time, the analytical study is different, because instead of a precision study, we will be working through an instrument carryover study. With this study, we will try to understand the differences when using the same analyte and the same type of specimen, but in a different type of analytical study.

In a carryover study, we would investigate the effect of possible carryover of materials from positive samples to a following negative sample. Let's discuss what carryover is and how it can occur. We have an instrument and a sample that is a true negative. This sample is surrounded by very high-positive samples, the highest possible value in the intended use population. We would like to see whether these very high-positive samples surrounding the negative sample will affect the negative sample; we want to see if there is a carryover effect that causes the negative sample to have a positive result.

A carryover study requires a true-negative sample. We also need to have at least five runs, each with a negative sample surrounded by one or more high-positive samples. This means that we need to have a lot of replicates of the same negative sample, so we need to have enough volume of the negative sample in order to conduct the carryover study.

G2 Principles

There are seven principles applicable to this case study.

G2.1 Principle 1 - Identify the Objective for Using Surrogate Samples for This Study

A sample from an individual patient would not provide enough sample volume to do at least five runs and a few replicates for each run.

Appendix G. (Continued)

G2.2 Principle 2 - Identify Critical Factors Related to the Performance Study Type

We need to have a sequence of high-positive samples and a negative sample. All systems are different and there can be some nuances, so during your risk assessment, you need to understand what kinds of additional factors are important for the carryover study. These can be issues related to the type of sample that is used for analysis, for example, before or after extraction and before or after the washing process.

Another example could be that, in a particular corner of a plate, the negative sample will be surrounded not by six samples or by four, but maybe by fewer samples. Such nuances are important for design of the carryover study.

G2.3 Principle 3 - Consider Factors Based on Whether the Test Is Qualitative, Quantitative, or Semi-Quantitative

This is a qualitative test, so we need to think about what that means for our carryover study design. For a qualitative carryover assessment, we will need to have a sequence of high-positive samples and the negative sample.

G2.4 Principle 4 - Consider the Effect of Sample Processing Such as Extraction Method, If Applicable

Bacterial lysis and nucleic acid extraction of the specimen could result in the presence of naturally occurring interfering materials. Though this is not anticipated, it is also not predictable; thus, it is something to be aware of for this study. If the extraction method could influence the patient sample, then it should be appropriately modeled in the carryover study.

G2.5 Principle 5 - Evaluate Property Characteristics of the Matrix

Remember that for a carryover study, we will alternate negative and high-positive samples. If we were to see that the negative sample became positive, we would conclude that it is due to a carryover effect. As we have already discussed, we cannot be completely certain that a patient sample is a true negative with chlamydia testing, and we need a true-negative sample for a carryover study.

A common scenario demonstrates the difficulty in obtaining true-negative patient samples: perhaps you have received residual patient samples that were tested for the analyte of interest by some other test and that test result was negative. Can you be sure, based on the other test result, that the residual sample is truly a negative sample? Likely not. As every test has some limit of detection, there is still some possibility that this sample is not truly negative. There is a chance that if the negative sample is tested again, it will be a positive result, not because of carryover, but rather because the sample is not a true negative. So, the sample must be a true negative. When we prepare the true-negative sample, naturally present organisms different from the analyte of interest or its extract in the matrix should be included to simulate potential interferences.

Appendix G. (Continued)

G2.6 Principle 6 - Evaluate Property Characteristics of the Analyte

Remember that for a carryover study, we need very, very high-positive samples. Generally, it is difficult to find very high-positive patient samples, though some exist. Therefore, it is acceptable to use biological analyte, such as cell culture or a nucleic acid sequence of chlamydia. A true-negative sample should have negative matrix with added human cells and other microorganisms that we routinely observe in vaginal or cervical swabs.

G2.7 Principle 7 - Assess Analyte and Matrix Combinations of Surrogate Samples

The need to assess whether there is a synergistic interaction in the combination of analyte and a particular matrix is necessary. An analyte that mimics the biological analyte is needed, and the analyte concentration needs to be controlled to make sure that negative samples are true negatives. Cell culture is done for chlamydia to determine if spiking is needed in the CT material. Again, the primary concern is to make sure that the sample is truly negative.

When run in a molecular test, samples that tested negative in cell culture may produce a low level of analyte signal that could be mistaken for carryover. This is likely not a carryover effect, but rather due to the exquisite analytical sensitivity of molecular tests. Therefore, the use of surrogates for matrix in molecular analysis is justifiable for the purposes of a carryover study.

G3 Navigating the Hierarchy

Here, the hierarchy for an instrument carryover study where the analyte is chlamydia is navigated. The specimen is a vaginal or endocervical swab. For the carryover study, two types of samples are needed: a true negative that definitively has no analyte, and a very, very high-positive sample.

Case Study 3

Hierarchy: CT - Endocervical or vaginal swab

Instrument Carryover
A1
A2
B1
B2
C1
C2
D1
D2
E1
E2
F1
F2
G
H

A1

Analyte: Unspiked, Individual

Matrix: Biological, Unaltered

Example Endocervical or vaginal swab

Challenges

- Inherent patient sample matrix variability inhibits the characterization of factors that contribute to carryover
- TN - Difficult to isolate and identify “residual” materials from patient sample that potentially affect carryover
- TN & HP - Insufficient sample volume for size of study

Outcome

Go to the next level

Figure G1. Analyte: Unspiked, Individual; Matrix: Unaltered, Biological

Appendix G. (Continued)

A1 is a sample from an individual patient. The matrix is biological and unaltered. An example of an A1 surrogate sample is an endocervical or vaginal swab from one patient. This sample cannot be used as a true-negative sample, due to the concerns noted previously. The sample is unspiked and is unlikely to be acceptable as a very high positive. In addition, the volume of a single patient sample will not be sufficient for carryover studies.

Case Study 3

Hierarchy: CT - Endocervical or vaginal swab

Instrument Carryover
A1
A2
B1
B2
C1
C2
D1
D2
E1
E2
F1
F2
G
H

A2

Analyte:
Unspiked, Pooled

Matrix:
Biological, Unaltered

Example
Individual extracted endocervical or vaginal swab combined into a sample

Challenges

- HP - Pool may “dilute” impact of the potential cause of carryover
- TN - Pool might not be a true negative
- HP - Difficult to evaluate one organism if method is designed for multiple organisms

Outcome

[Go to the next level](#)

Figure G2. Analyte: Unspiked, Pooled; Matrix: Unaltered, Biological

In A2, the surrogate sample is a pooled sample from more than one patient. The matrix is biological and unaltered. An example of this A2 surrogate is a pooled swab sample from multiple patients. A2 solves the volume problem, but we still do not know if this pooled sample is a true negative. For the very high-positive sample, A2 solves the volume problem, but it is difficult to have very high concentrations. In addition, it can be difficult to evaluate one organism if the method is designed for multiple organisms.

Appendix G. (Continued)

Case Study 3

Hierarchy: CT - Endocervical or vaginal swab

Instrument Carryover
A1
A2
B1
B2
C1
C2
D1
D2
E1
E2
F1
F2
G
H

B1

Analyte:

Spiked - Biological, Individual

Matrix:

Biological, Unaltered

Example

Endocervical or vaginal swab spiked with biological CT organism

Challenges

- TN - N/A, because a negative should not be supplemented with analyte
- HP - Inadequate volume from an individual patient sample and insufficient sample volume for the size of the study
- HP - Characteristics that might impact carryover are varied and undefined, so an individual patient sample is not characteristic of the entire population tested

Outcome

[Go to the next level](#)

Figure G3. Analyte: Spiked - Biological, Individual; Matrix: Unaltered, Biological

Here, we have B1. The sample is from an individual patient, and the sample is spiked with biological analyte. The matrix is biological and unaltered. An example of this B1 surrogate is an endocervical or vaginal swab from an individual patient, and it is spiked with biological CT. We've already discussed the issue for the true-negative sample.

For the high positive, individual patient samples present volume limitations, and there is some chance that this sample can be atypical (may not be representative of the entire population tested).

Appendix G. (Continued)

Case Study 3

Hierarchy: CT - Endocervical or vaginal swab

Instrument Carryover
A1
A2
B1
B2
C1
C2
D1
D2
E1
E2
F1
F2
G
H

B2

Analyte:

Spiked - Biological, Pooled

Matrix:

Biological, Unaltered

Example

Pool of individual extracted endocervical or vaginal swabs spiked with biological CT organism

Challenges

- TN - N/A, because a negative should not be supplemented with analyte
- HP - Pool may “dilute” the impact of the potential cause of carryover
- HP - Difficult to evaluate one organism if method is designed for multiple organisms

Outcome

[Go to the next level](#)

Figure G4. Analyte: Spiked - Biological, Pooled; Matrix: Unaltered, Biological

B2 is next. The sample is pooled from more than one patient, and this pool of samples is spiked with a biological analyte. The matrix is biological and unaltered. An example of this B2 surrogate is a pool of individual endocervical or vaginal swabs spiked with biological CT organism. We encounter similar challenges for B2 as for the previous options. For the high positive, we need to understand whether pooling can change the matrix in such a way that it will be more difficult to detect the potential cause of carryover, so it is important to know the properties of the sample tested on the instrument. For a CT assay with vaginal swabs, we know that the swab is put in transport media, so we do not anticipate that combining two transport media with patient samples will drastically change the matrix of the pooled sample. Pooling of the sample can also ensure that any potential atypical individual samples do not affect the results and that the carryover study reflects typical real-life performance. If the assay detects multiple organisms, then positive samples should reflect very high values for each of the detected organisms. This can be a very challenging sample, and at times, B2 surrogate samples can be helpful.

Appendix G. (Continued)

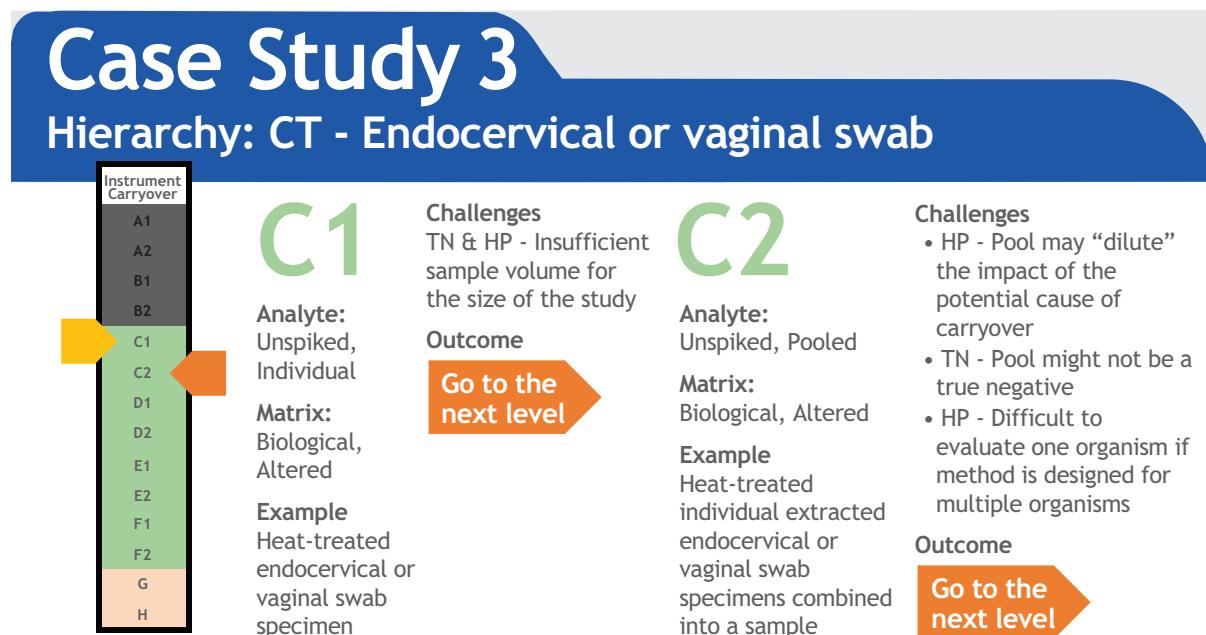


Figure G5. Analyte: Unspiked, Individual or Pooled; Matrix: Altered, Biological

C samples are where the sample is unspiked and the matrix is altered. The C1 sample is an individual patient sample. The matrix is biological and altered. An example of a C1 surrogate sample is a heat-treated endocervical or vaginal swab specimen. Volume limitations are a challenge if we use C1.

The C2 sample is pooled from more than one patient. The matrix is biological and altered. An example of a C2 surrogate sample is heat-treated individual endocervical or vaginal swab specimens combined into a single sample. With C2, we know that the sample is a true negative. But the matrix can be so altered that it does not represent patient samples. We know that an approach of treating the sample in order to have zero concentration is not used for chlamydia analyte.

Appendix G. (Continued)

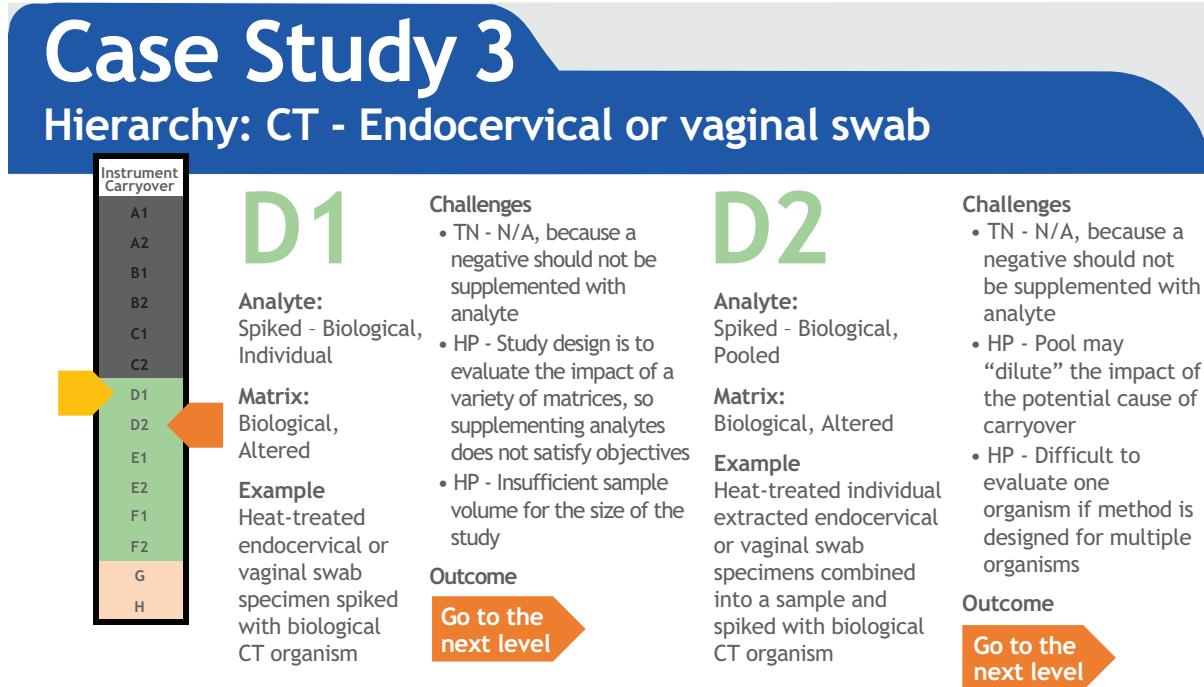


Figure G6. Analyte: Spiked - Biological, Individual or Pooled; Matrix: Altered, Biological

D are samples in which the matrix is still biological, but it is altered, and the sample was spiked with a biological analyte. The D1 analyte is an individual patient sample spiked with biological analyte, and the D2 analyte is pooled from more than one patient and spiked with biological analyte. The matrix for D1 and D2 is an altered biological matrix.

For our true-negative sample, neither D1 nor D2 is appropriate, because a negative sample should not be supplemented with analyte. For our high-positive sample, neither D1 nor D2 is appropriate, because treating the sample for zero concentration is not a practical approach for chlamydia analyte.

Appendix G. (Continued)

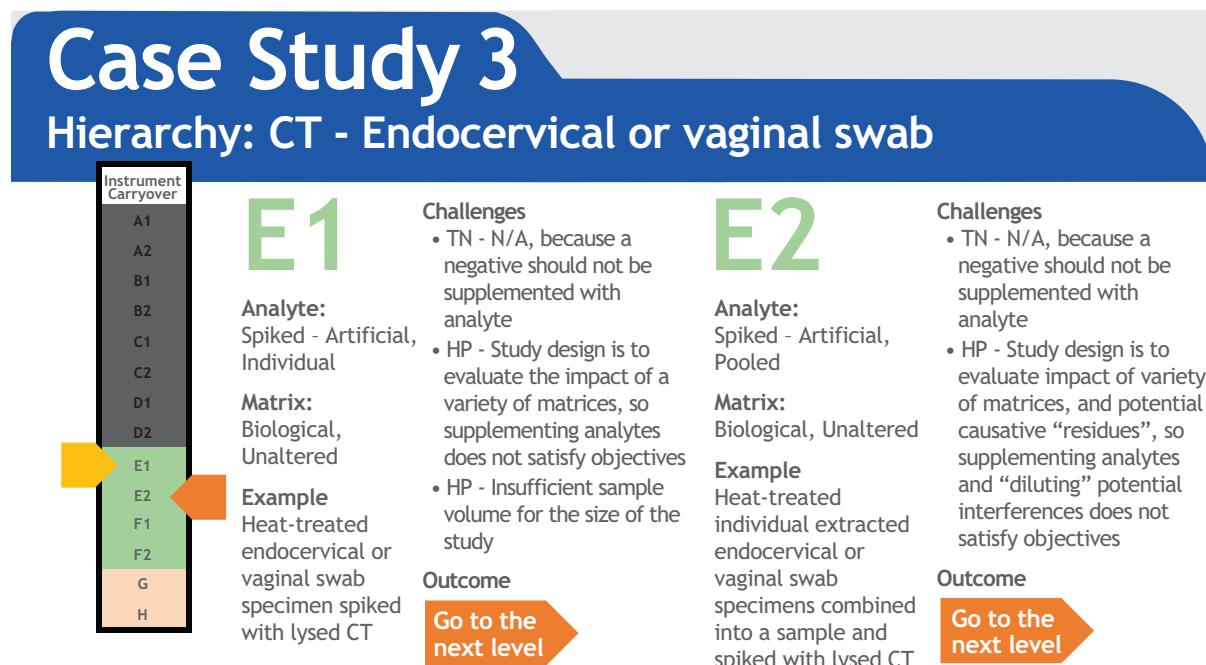


Figure G7. Analyte: Spiked - Artificial, Individual or Pooled; Matrix: Unaltered, Biological

Now we're moving on to the E samples, where the sample is spiked with artificial analyte.

In E1, the sample is from an individual patient and spiked with an artificial analyte. The matrix is biological and unaltered. An example of an E1 surrogate sample is a single patient swab specimen spiked with CT cell lysate. The challenge with E1 is that we may have volume limitations that prevent us from performing a large study. As we've seen before, we can't use this sample as a negative, because we wouldn't supplement a negative with analyte.

In E2, the sample is pooled from more than one patient and spiked with an artificial analyte. The matrix is biological and unaltered. An example of an E2 surrogate sample is individual swab specimens combined into a sample and spiked with CT cell lysate. We run into the same challenges as E1 regarding the negative. For the HP, E2 is also not appropriate, because the supplemented analyte is artificial.

Appendix G. (Continued)

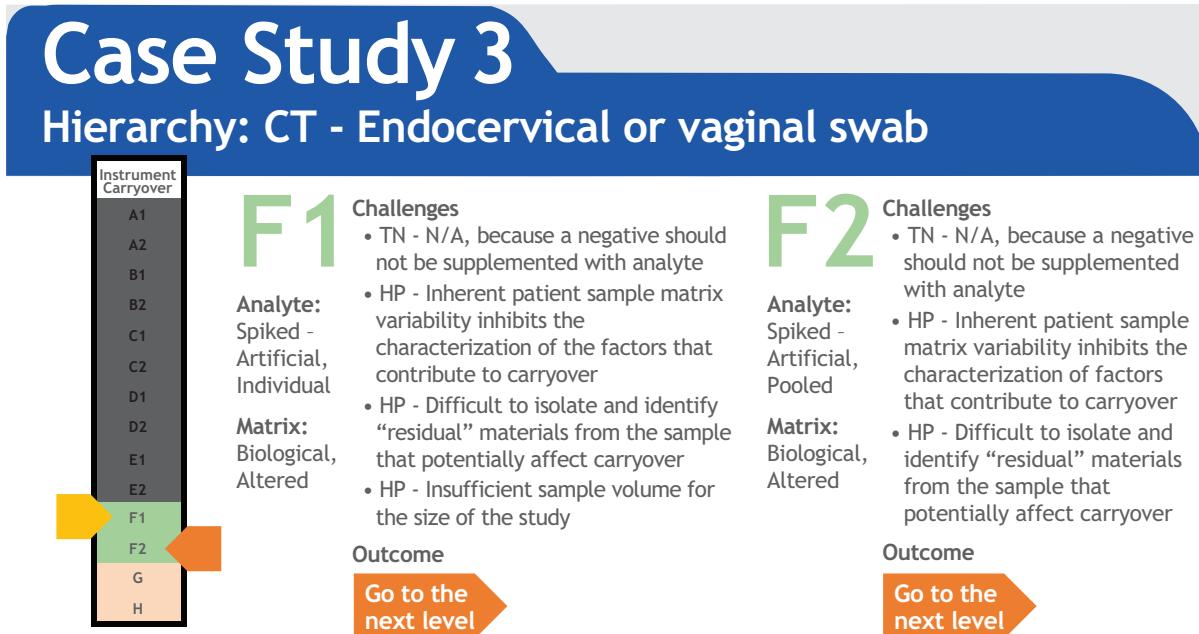


Figure G8. Analyte: Spiked - Artificial, Individual or Pooled; Matrix: Altered, Biological

In the F samples, we still have a biological, but altered, matrix. The sample is spiked, but this time with an artificial analyte instead of a biological analyte. The main challenge with both F1 and F2 is that it may be difficult to isolate the chlamydia analyte in order to have samples with zero concentration. So, altered biological matrix can be a difficult scenario for the chlamydia assays. Move to the next level.

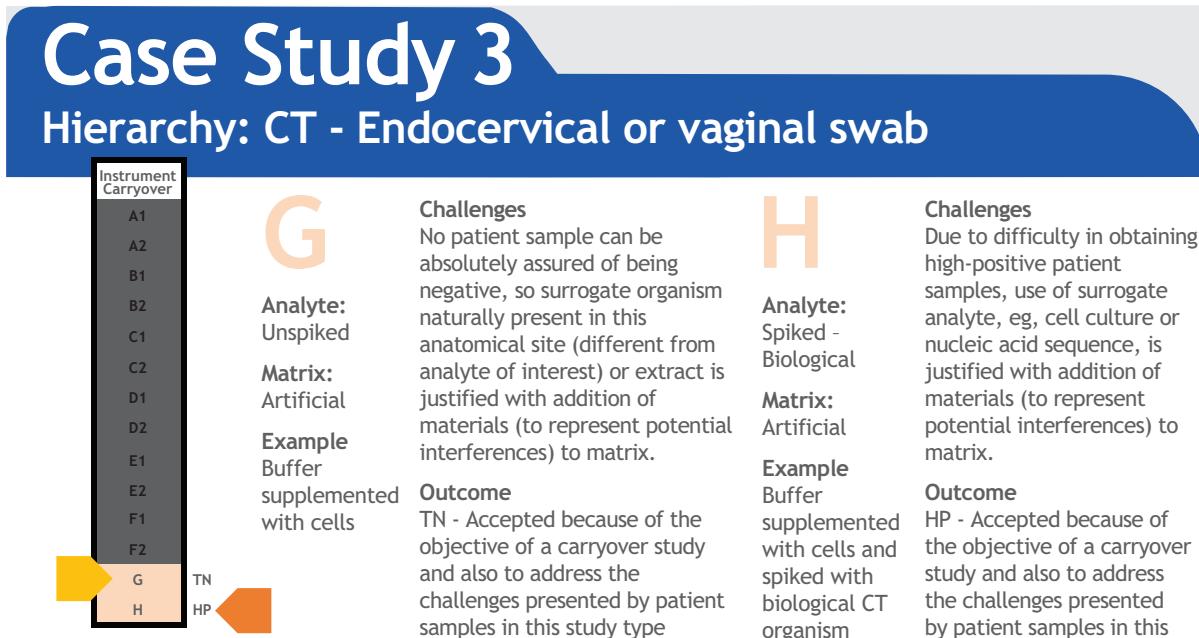


Figure G9. Analyte: Unspiked or Spiked - Biological; Matrix: Artificial

Appendix G. (Continued)

Samples G and H both have an artificial matrix. G is not spiked with any analyte. It is a completely negative matrix. An example of surrogate sample G would be a buffer containing cells that are found in a vaginal swab. Surrogate sample G has no volume limitations and certainly is completely negative. Therefore, it is appropriate for the true negative. A note on the hierarchy next to G has been added.

H is spiked with a biological analyte. An example of surrogate sample H would be buffer supplemented with cells that are found in typical vaginal swabs and spiked with biological CT organism. Remember that purified CT obtained from cell culture is a biological analyte.

G4 Conclusion

Case Study 3 Conclusion

*Chlamydia trachomatis: Instrument Carryover Study
Specimen: Vaginal or Endocervical Swab*

Surrogate Sample	Description
High positive (HP)	H Analyte: Spiked - Biological Matrix: Artificial
True negative (TN)	G Analyte: Unspiked Matrix: Artificial

Figure G10. Case Study 3 - Conclusion

Surrogate sample H is acceptable for the HP, because we can control the analyte concentration and get as high as we need to for the carryover study. We have made that note on the hierarchy in Figure G10.

To recap for this case, a CT carryover study using vaginal or endocervical swab specimens, two surrogate sample types were needed: a high positive and true negative.

For the high positive, B2 (a pool of patient samples spiked with biological analyte) was the best choice. Due to the high volume requirement and logistical issues, it is acceptable to use artificial matrix instead of patient samples, and then spike with biological analyte. This is surrogate sample H, which is a surrogate sample G matrix spiked with biological analyte. For the true negative, we arrived at surrogate sample G, an artificial matrix supplemented with cells found in the typical vaginal swabs.

The Quality Management System Approach

Clinical and Laboratory Standards Institute (CLSI) subscribes to a quality management system (QMS) approach in the development of standards and guidelines that facilitates project management, defines a document structure using a template, and provides a process to identify needed documents. The QMS 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:

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

The QSEs covered by EP39 and its related CLSI documents are available on the CLSI website: <https://clsi.org/qse>

Related CLSI Reference Materials^a

- C37** **Preparation and Validation of Commutable Frozen Human Serum Pools as Secondary Reference Materials for Cholesterol Measurement Procedures. 1st ed., 1999.** This guideline details procedures for the manufacture and evaluation of human serum pools for cholesterol measurement.
- C49** **Analysis of Body Fluids in Clinical Chemistry. 2nd ed., 2018.** This guideline provides information for the medical laboratory for evaluating measurement procedures, as well as a strategy to characterize assay performance, when applied to body fluid matrixes. Key concepts that apply to the entire test cycle, including preexamination, examination, and postexamination phases of body fluid testing, are discussed.
- EP05** **Evaluation of Precision of Quantitative Measurement Procedures. 3rd ed., 2014.** This document provides guidance for evaluating the precision performance of quantitative measurement procedures. It is intended for manufacturers of quantitative measurement procedures and for laboratories that develop or modify such procedures.
- EP06** **Evaluation of Linearity of Quantitative Measurement Procedures. 2nd ed., 2020.** This guideline provides information for characterizing the linearity interval of a measurement procedure, validating a linearity interval claim (to be performed by the manufacturer), and verifying an established linearity interval claim (to be performed by the end user).
- EP07** **Interference Testing in Clinical Chemistry. 3rd ed., 2018.** This guideline provides background information, guidance, and experimental procedures for investigating, identifying, and characterizing the effects of interferents on clinical chemistry test results.
- EP09** **Measurement Procedure Comparison and Bias Estimation Using Patient Samples. 3rd ed., 2018.** 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.
- EP12** **User Protocol for Evaluation of Qualitative Test Performance. 2nd ed., 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** **Evaluation of Commutability of Processed Samples. 3rd ed., 2014.** This document provides guidance for evaluating the commutability of processed samples by determining if they behave differently than unprocessed patient samples when two quantitative measurement procedures are compared.

^a 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)

- EP17 Evaluation of Detection Capability for Clinical Laboratory Measurement Procedures. 2nd ed., 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.
- EP18 Risk Management Techniques to Identify and Control Laboratory Error Sources. 2nd ed., 2009.** This guideline describes risk management techniques that will aid in identifying, understanding, and managing sources of failure (potential failure modes) and help to ensure correct results. Although intended primarily for *in vitro* diagnostics, this document will also serve as a reference for clinical laboratory managers and supervisors who wish to learn about risk management techniques and processes.
- EP19 A Framework for Using CLSI Documents to Evaluate Clinical Laboratory Measurement Procedures. 2nd ed., 2015.** This report uses the "measurement procedure lifecycle" framework to aid users of CLSI evaluation protocols documents during establishment and implementation of measurement procedures developed by both commercial manufacturers and clinical laboratories, ie, for laboratory-developed tests.
- EP23™ Laboratory Quality Control Based on Risk Management. 1st ed., 2011.** This document provides guidance based on risk management for laboratories to develop quality control plans tailored to the particular combination of measuring system, laboratory setting, and clinical application of the test.
- EP25 Evaluation of Stability of *In Vitro* Diagnostic Reagents. 1st ed., 2009.** This document provides guidance for establishing shelf-life and in-use stability claims for *in vitro* diagnostic reagents such as reagent kits, calibrators, and control products.
- EP30 Characterization and Qualification of Commutable Reference Materials for Laboratory Medicine. 1st ed., 2010.** This document provides information to help material manufacturers in the production and characterization of commutable reference materials, as well as to assist assay manufacturers and laboratorians in the appropriate use of these materials for calibration and trueness assessment of *in vitro* diagnostic medical devices.
- EP34 Establishing and Verifying an Extended Measuring Interval Through Specimen Dilution and Spiking. 1st ed., 2018.** It is often medically necessary to provide results for specimens with concentrations above the analytical measuring interval of an *in vitro* diagnostic measurement procedure. This guideline helps manufacturers and laboratory scientists with establishing, validating, or verifying a dilution scheme that will provide an extended measuring interval for such specimens.
- EP35 Assessment of Equivalence or Suitability of Specimen Types for Medical Laboratory Measurement Procedures. 1st ed., 2019.** This guideline provides recommendations for assessing clinically equivalent performance for additional similar-matrix specimen types and suitable performance for dissimilar-matrix specimen types, such that the laboratory does not necessarily need to repeat the full measurement procedure validation for each specimen type. The recommendations in this guideline apply to both quantitative measurement procedures and qualitative examinations.

Related CLSI Reference Materials (Continued)

- EP37** **Supplemental Tables for Interference Testing in Clinical Chemistry. 1st ed., 2018.** This document includes recommended testing concentrations for analytes and endogenous substances that may interfere in clinical chemistry measurement procedures and is intended for use with the evaluation procedures in the Clinical and Laboratory Standards Institute guideline EP07.
- M29** **Protection of Laboratory Workers From Occupationally Acquired Infections. 4th ed., 2014.** 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.
- MM03** **Molecular Diagnostic Methods for Infectious Diseases. 3rd ed., 2015.** This report addresses topics relating to clinical applications, amplified and nonamplified nucleic acid methods, selection and qualification of nucleic acid sequences, establishment and evaluation of test performance characteristics, inhibitors, and interfering substances, controlling false-positive reactions, reporting and interpretation of results, quality assurance, regulatory issues, and recommendations for manufacturers and clinical laboratories.
- MM06** **Quantitative Molecular Methods for Infectious Diseases. 2nd ed., 2010.** This document provides guidance for the development and use of quantitative molecular methods, such as nucleic acid probes and nucleic acid amplification techniques of the target sequences specific to particular microorganisms. It also presents recommendations for quality assurance, proficiency testing, and interpretation of results.
- MM17** **Validation and Verification of Multiplex Nucleic Acid Assays. 2nd ed., 2018.** This guideline includes recommendations for analytical validation and verification of multiplex assays, as well as a review of different types of biological and synthetic reference materials.

NOTES

Discover How CLSI Can Improve Your Organization



The leading source for the latest medical laboratory standards.



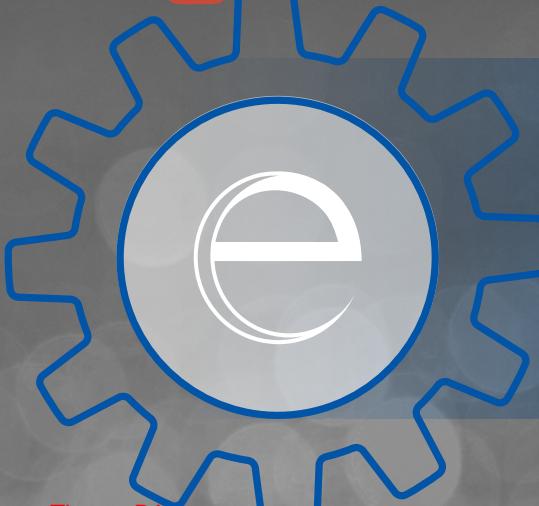
CLSI membership lets you directly impact best practice standards used to improve patient care worldwide—standards you use every day. Membership provides you with standards access, volunteering opportunities, influence in the standards development process, networking opportunities, discounts, and more.

Discover the membership option for you at clsi.org/join.



Our educational and training programs provide convenient, cost-effective continuing education and training resources to help you advance your professional development. We have a variety of easy-to-use, online educational resources and in-person trainings that make learning stress-free and convenient for you and your staff.

See our current offerings at clsi.org/global-training.



Ensure high-quality laboratory testing with CLSI standards. eCLIPSE Ultimate Access™, our complete online library of standards, makes it easy for you and your staff to quickly find the CLSI resources you need. Read, search, link, annotate, bookmark, and share notes with your staff, all within one easy-to-use platform.

Learn more at clsi.org/eCLIPSE.



CLINICAL AND
LABORATORY
STANDARDS
INSTITUTE®

PRINT ISBN 978-1-68440-130-7

ELECTRONIC ISBN 978-1-68440-131-4

EP39-Ed1