C28-A3c
Vol. 28 No. 30
Replaces C28-A2
Vol. 20 No. 13

Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline—Third Edition

This document contains guidelines for determining reference values and reference intervals for quantitative clinical laboratory tests.

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

Clinical and Laboratory Standards Institute

Advancing Quality in Health Care Testing

Clinical and Laboratory Standards Institute (CLSI, formerly NCCLS) is an international, interdisciplinary, nonprofit, standards-developing, and educational organization that promotes the development and use of voluntary consensus standards and guidelines within the health care community. It is recognized worldwide for the application of its unique consensus process in the development of standards and guidelines for patient testing and related health care issues. Our process is based on the principle that consensus is an effective and cost-effective way to improve patient testing and health care services.

In addition to developing and promoting the use of voluntary consensus standards and guidelines, we provide an open and unbiased forum to address critical issues affecting the quality of patient testing and health care.

PUBLICATIONS

A document is published as a standard, guideline, or committee report.

Standard A document developed through the consensus process that clearly identifies specific, essential requirements for materials, methods, or practices for use in an unmodified form. A standard may, in addition, contain discretionary elements, which are clearly identified.

Guideline A document developed through the consensus process describing criteria for a general operating practice, procedure, or material for voluntary use. A guideline may be used as written or modified by the user to fit specific needs.

Report A document that has not been subjected to consensus review and is released by the Board of Directors.

CONSENSUS PROCESS

The CLSI voluntary consensus process is a protocol establishing formal criteria for:

- the authorization of a project
- the development and open review of documents
- the revision of documents in response to comments by users
- the acceptance of a document as a consensus standard or guideline.

Most documents are subject to two levels of consensus—"proposed" and "approved." Depending on the need for field evaluation or data collection, documents may also be made available for review at an intermediate consensus level.

Proposed A consensus document undergoes the first stage of review by the health care community as a proposed standard or guideline. The document should receive a wide and thorough technical review, including an overall review of its scope, approach, and utility, and a line-by-line review of its technical and editorial content.

Approved An approved standard or guideline has achieved consensus within the health care community. It should be reviewed to assess the utility of the final document, to ensure attainment of consensus (ie, that comments on earlier versions have been satisfactorily addressed), and to identify the need for additional consensus documents.

Our standards and guidelines represent a consensus opinion on good practices and reflect the substantial agreement by materially affected, competent, and interested parties obtained by following CLSI's established consensus procedures. Provisions in CLSI standards and guidelines may be more or less stringent than applicable regulations. Consequently, conformance to this voluntary consensus document does not relieve the user of responsibility for compliance with applicable regulations.

COMMENTS

The comments of users are essential to the consensus process. Anyone may submit a comment, and all comments are addressed, according to the consensus process, by the committee that wrote the document. All comments, including those that result in a change to the document when published at the next consensus level and those that do not result in a change, are responded to by the committee in an appendix to the document. Readers are strongly encouraged to comment in any form and at any time on any document. Address comments to Clinical and Laboratory Standards Institute, 940 West Valley Road, Suite 1400, Wayne, PA 19087, USA.

VOLUNTEER PARTICIPATION

Health care professionals in all specialties are urged to volunteer for participation in CLSI projects. Please contact us at customerservice@clsi.org or +610.688.0100 for additional information on committee participation.

C28-A3c
ISBN 1-56238-682-4
Volume 28 Number 30
ISSN 0273-3099

Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline—Third Edition

Gary L. Horowitz, MD, Chairholder
Sousan Altaie, PhD
James C. Boyd, MD
Ferruccio Ceriotti, MD
Uttam Garg, PhD, DABCC
Paul Horn, PhD
Amadeo Pesce, PhD
Harrison E. Sine, PhD
Jack Zakowski, PhD, FACB

Abstract

Clinical and Laboratory Standards Institute document C28-A3c—*Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline—Third Edition* is written for users of diagnostic laboratory tests. It offers a protocol for determining reference intervals that meet the minimum requirements for reliability and usefulness. The guideline focuses on health-associated reference values as they relate to quantitative clinical laboratory tests. Included are various requirements for studies to determine reference values for a new analyte or a new analytical method of a previously measured analyte. Also discussed is the transfer of established reference values from one laboratory to another.

Clinical and Laboratory Standards Institute (CLSI). *Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline—Third Edition* CLSI document C28-A3c (ISBN 1-56238-682-4). Clinical and Laboratory Standards Institute, 940 West Valley Road, Suite 1400, Wayne, Pennsylvania 19087-1898 USA, 2008.

Suggested Citation

(CLSI. *Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline—Third Edition.* CLSI document C28-A3c. Wayne, PA: Clinical and Laboratory Standards Institute; 2008.)

Proposed Guideline
March 1992

Approved Guideline
June 1995

Approved Guideline—Second Edition
June 2000

Proposed Guideline—Third Edition
March 2008

Approved Guideline—Third Edition
November 2008
October 2010 (corrected version)

ISBN 1-56238-682-4
ISSN 0273-3099

Committee Membership

Area Committee on Clinical Chemistry and Toxicology

David A. Armbruster, PhD, DABCC, FACB
Chairholder
Abbott
Abbott Park, Illinois

Christopher M. Lehman, MD
Vice-Chairholder
Univ. of Utah Health Sciences Center
Salt Lake City, Utah

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

David M. Bunk, PhD
National Institute of Standards and Technology
Gaithersburg, Maryland

David G. Grenache, PhD, MT(ASCP), DABCC
University of Utah, ARUP Laboratories
Salt Lake City, Utah

Steven C. Kazmierczak, PhD, DABCC, FACB
Oregon Health and Science University
Portland, Oregon

Linda Thienpont, PhD
University of Ghent
Ghent, Belgium

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

Hubert Vesper, PhD
Centers for Disease Control and Prevention
Atlanta, Georgia

Jack Zakowski, PhD, FACB
Beckman Coulter, Inc.
Brea, California

Advisors

Mary F. Burritt, PhD
Mayo Clinic
Scottsdale, Arizona

Paul D'Orazio, PhD
Instrumentation Laboratory
Lexington, Massachusetts

Carl C. Garber, PhD, FACB
Quest Diagnostics, Incorporated
Lyndhurst, New Jersey

Uttam Garg, PhD, DABCC
Children's Mercy Hospital and Clinic
Kansas City, Missouri

Neil Greenberg, PhD
Ortho-Clinical Diagnostics, Inc.
Rochester, New York

Harvey W. Kaufman, MD
Quest Diagnostics, Inc.
Lyndhurst, New Jersey

W. Gregory Miller, PhD
Virginia Commonwealth University
Richmond, Virginia

Gary L. Myers, PhD
Centers for Disease Control and Prevention
Atlanta, Georgia

David Sacks, MD
Brigham and Women's Hospital and Harvard Medical School
Boston, Massachusetts

Bette Seamonds, PhD
Mercy Health Laboratory
Swarthmore, Pennsylvania

Dietmar Stöckl, PhD
STT Consulting
Horebeke, Belgium

Thomas L. Williams, MD
Nebraska Methodist Hospital
Omaha, Nebraska

Working Group on Reference Intervals

Gary L. Horowitz, MD
Chairholder
Beth Israel Deaconess Medical Center
Boston, Massachusetts

Sousan S. Altaie, PhD
FDA Ctr. for Devices/Rad. Health
Rockville, Maryland

James C. Boyd, MD
UVA Health System
Charlottesville, Virginia

Ferruccio Ceriotti, MD
Diagnostica E Ricerca San Raffaele
Milano, Italy

Uttam Garg, PhD, DABCC
Children's Mercy Hospitals and Clinics
Kansas City, Missouri

Amadeo Pesce, PhD
University of Cincinnati College of Medicine
Cincinnati, Ohio

Harrison E. Sine, PhD
Roche Diagnostics, Inc.
Indianapolis, Indiana

Jack Zakowski, PhD, FACB
Beckman Coulter, Inc.
Brea, California

Advisors

Paul S. Horn, PhD
University of Cincinnati
Psychiatry Service, Veterans Affairs Medical Center Cincinnati
Cincinnati, Ohio

James J. Miller, PhD, DABCC, FACB
University of Louisville School of Medicine
Louisville, Kentucky

Staff

Clinical and Laboratory Standards Institute
Wayne, Pennsylvania

Lois M. Schmidt, DA
Vice President, Standards Development and Marketing

Jane M. Oates, BS, MT(ASCP)
Staff Liaison

Ron S. Quicho
Project Manager

Melissa A. Lewis
Editor

Acknowledgment

This guideline was prepared by Clinical and Laboratory Standards Institute (CLSI), as part of a cooperative effort with IFCC to work toward the advancement and dissemination of laboratory standards on a worldwide basis. CLSI gratefully acknowledges the participation of IFCC in this project. The IFCC experts for this project are Ferruccio Ceriotti, MD, Diagnostica E Ricerca San Raffaele; and James C. Boyd, MD, UVA Health System.

Contents

Contents (Continued)

Foreword

A measured or observed laboratory test result from a person (usually a patient) is compared with a reference interval for the purpose of making a medical diagnosis, therapeutic management decision, or other physiological assessment. The interpretation of clinical laboratory data is, therefore, a comparative decision-making process. For this decision-making process to occur, reference values are needed for all tests in the clinical laboratory, and the provision of reliable reference intervals is an important task for clinical laboratories and diagnostic test manufacturers. The reference values most commonly used (known as "normal values" and sometimes "expected values") have traditionally been poorly defined and certainly not determined by a uniform process. It is now apparent that it is important to develop reference intervals using a more systematic process that takes into account the various influences on the measured laboratory test results.

A theory of reference values that provides definitions, principles, and procedures for the determination and use of reference values was developed by the Expert Panel on Theory of Reference Values (EPTRV) of the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) and the Standing Committee on Reference Values of the International Council for Standardization in Haematology (ICSH). The fruits of the tireless labors of these committees appear in a series of articles[1-6] that provide a rational approach and sound basis for the determination of reference values. These definitions also provided a basis for the development of this guideline. CLSI is indebted to the members of the IFCC committee and to the many other investigators who contributed to this discipline and upon whose knowledge it has drawn.

This guideline begins with definitions proposed by the EPTRV of the IFCC that are important to the discussion of reference values. An outline of the broad procedural protocol for establishing reference intervals is included, followed by specifics of each of the composite processes. Issues related to the reference subject selection process, the importance of preanalytical and analytical considerations, the calculation methods and requirements for estimating valid reference intervals, and the transference of reference intervals are discussed. Examples of the recommended estimation and calculation processes are provided. Finally, issues related to the presentation and use of reference intervals are discussed, followed by a brief section that examines a number of important but collateral reference value topics not amenable to inclusion in this document.

Key Words

Critical value, observed value, reference distribution, reference individual, reference interval, reference limit, reference population, reference sample group, reference value

Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline—Third Edition

1 Scope

This document provides diagnostic laboratories and diagnostic test manufacturers with updated guidelines for determining reference intervals for quantitative laboratory tests. It includes specific recommendations regarding procedures that can be used to establish and verify reliable reference intervals for use in clinical laboratory medicine. By following these recommendations, laboratories can provide reference intervals that are adequate and useful for clinical interpretation.

Issues related to the reference subject selection process, the importance of preanalytical and analytical considerations, the calculation methods and requirements for estimating valid reference intervals, and the transference of reference intervals are discussed. Examples of the recommended estimation and calculation processes are provided. Finally, issues related to the presentation and use of reference intervals are discussed, followed by a brief section that examines a number of important but collateral reference value topics not amenable to inclusion in this document.

2 Introduction

Since the last update to this document (2000), two notable trends have emerged in clinical laboratory practice to which the working group would like to call attention.

First, for some analytes, reference intervals have been replaced by ***decision limits***, established by national (or international) consensus. As examples, consider cholesterol and glycated hemoglobin. For such analytes, there is no need to establish *de novo,* or even to verify, the reference intervals. Rather, laboratories must concern themselves with the accuracy of the results they report; that is, that cholesterol values they report are not appreciably different from the values that are reported by a certified reference laboratory on the same samples. For such analytes, the onus falls on manufacturers to ensure their methods are traceable (see CLSI document X05[7]) and on individual laboratories to ensure they run those methods correctly (using peer group quality control [QC], proficiency testing, etc.).

Second, the working group recognizes the reality that, in practice, very few laboratories perform their own reference interval studies. As indicated in this document, the working group endorses its previous recommendation that the ***best*** means to ***establish*** a reference interval is to collect samples from a sufficient number of qualified reference individuals to yield a minimum of 120 samples for analysis, by nonparametric means, for each partition (eg, sex, age range).

The fact of the matter, though, is that few laboratories, or even manufacturers, do such studies. Often, if any study is done, far fewer individuals are used, with assumptions made about the underlying distributions and about the comparability among partitions. Sometimes (eg, electrolytes), instead of performing a new reference interval study, laboratories and manufacturers refer to studies done many decades ago, when both the methods and the population were very different.

For these reasons, the working group believes strongly that individual laboratories should focus more on ***verifying*** reference intervals established elsewhere, a much less formidable task. As noted in this document, this can be done in at least two practical ways:

(1) If a laboratory has previously established a reference interval for its own population, then it can verify that reference interval by ***transference***, using a CLSI/NCCLS document EP09[8] protocol (see Section 10). ***A major advantage of this option is there is no need to collect samples from***

reference individuals. One can use existing patient samples, even from subjects not known to be healthy, thus overcoming one of the major obstacles in reference interval studies.

(2) As an alternative, a laboratory can verify a reference interval, established by more stringent techniques elsewhere, ***by collecting as few as 20 samples from qualified reference individuals.*** As noted in Section 11, with the data from these samples in hand, one can do a simple binomial test, or one can apply more sophisticated tests to achieve better sensitivity and specificity. Whichever method one chooses, though, the important point is, with as few as 20 samples from reference individuals, a laboratory can verify reasonably well the applicability of a reference interval to its own population and methodology.

The CLSI working group is encouraged by other developments that should make the establishment of reference intervals less formidable.

- The working group urges all manufacturers to ensure their methods exhibit traceability to appropriate standards (see CLSI document X05[7] and ISO 17511[9]) when they exist. As a result, values for many assays from different laboratories should be interchangeable, which may make it possible to combine data from multiple sites to establish reference intervals, thereby reducing the burden on each laboratory to collect samples from as many as 120 individuals (see Section 6.2).

- The working group calls attention to computer-intensive procedures that permit increased precision and less stringent sample size requirements to establish reference intervals. If a laboratory has adequate statistical and computing competence, the working group encourages consideration of procedures that do not require 120 individuals to estimate reference limits and confidence intervals (CI) (see Section 9).

In summary, the working group believes ***every laboratory is more than capable of verifying the applicability of reference intervals to its own population.*** In addition, the working group strongly endorses the recommendations of the previous working group on the proper way to establish a reference interval, and it extends those recommendations by introducing recommendations about multicenter studies and modern statistical methods.

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 feature of "universal precautions and body substance isolation" practices. Standard precautions cover the transmission of all infectious agents and thus are more comprehensive than universal precautions, which are intended to apply only to transmission of blood-borne pathogens. Standard and universal precaution guidelines are available from the US Centers for Disease Control and Prevention.[10] For specific precautions for preventing the laboratory transmission of all infectious agents from laboratory instruments and materials and for recommendations for the management of exposure to all infectious disease, refer to CLSI document M29.[11]

4 Terminology

4.1 A Note on Terminology

The document begins with the definition of certain terms that are important to the discussion of reference values. The terminology adopted is proposed by the Expert Panel on Theory of Reference Values (EPTRV) of the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC), which was carefully developed for a more systematic and unambiguous discussion. An outline of the broad

procedural protocol for establishing reference intervals is included, followed by specifics of each of the composite processes.

The following terms permit relatively unambiguous description and discussion of the subject of reference values. This list of definitions was proposed by the EPTRV of the IFCC[1] and International Council for Standardization in Haematology (ICSH), and was endorsed by the World Health Organization (WHO) and other organizations worldwide. These definitions represent what is becoming accepted universal terminology. A discussion and clarification of these terms follows (see Section 4.4).

4.2 Definitions

observed value (patient laboratory test result) – the value of a particular type of quantity, obtained by observation or measurement of a test subject (ie, patient), to be compared with reference values, reference distributions, reference limits, or reference intervals.

precision (of measurement) – closeness of agreement between independent test results obtained under stipulated conditions (ISO 3534-1).[12]

reference distribution – the distribution of reference values; **NOTE:** Hypotheses regarding the distribution of a reference population may be tested using the reference distribution of the reference sample group and adequate statistical methods. The parameters of the hypothetical distribution of the reference population may be estimated using the reference distribution of the reference sample group and adequate statistical methods.

reference individual – a person selected for testing on the basis of well-defined criteria; **NOTE:** It is usually important to define the person's state of health.

The following terms are used in connection with selecting reference individuals in the context of this document:

- ***a priori:*** application of criteria before the collection of samples.
- ***a posteriori:*** application of criteria after the collection of samples.

reference interval – the interval between, and including, two reference limits; **NOTE:** It is designated as the interval of values from the lower reference limit to the upper reference limit (eg, for calcium, the reference interval is 9.1 mg/dL to 10.3 mg/dL [2.27 mmol/L to 2.57 mmol/L]; in some cases, only one reference limit is important, usually an upper limit, "*x*," and the corresponding reference interval is 0 to *x*).

The following terms are used in connection with reference intervals in the context of this document:

- **defining a reference interval –** describing in detail the characteristics of the reference interval (ie, the central 95% of apparently healthy men and women between the ages of 18 and 65).
- **establishing (or determining) a reference interval** – the process used in creating a reference interval *de novo*, encompassing all of the steps from selection of reference individuals, through exact details of the analytical methods, and concluding with data collection and analysis.
- **transferring a reference interval –** the process by which one may be able to adapt a previously established reference interval to a new analytical method or to a new location.

- **verifying (or validating) a reference interval –** the process by which one ensures, with reasonable confidence, using a relatively small number of reference individuals (eg, n = 20), that a reference interval **established** elsewhere, or **transferred** from another study, can be used locally.

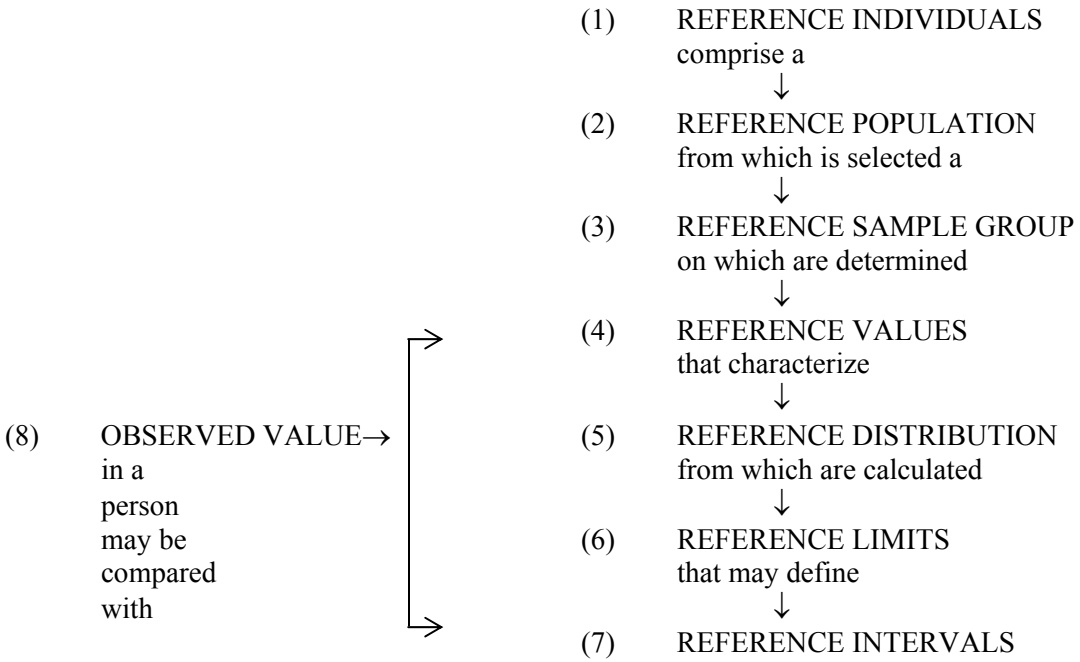
reference limit – a value derived from the reference distribution and used for descriptive purposes; **NOTE:** It is common practice to define a reference limit so a stated fraction of the reference values is less than or equal, or greater than or equal, to the respective upper or lower limit; the reference limit is descriptive of the reference values and may be distinguished from various other types of decision limits.

reference population – a group consisting of all the reference individuals; **NOTE 1:** The reference population usually has an unknown number of members and, therefore, is a hypothetical entity; the reference population may consist of only one member (eg, a person may serve as a reference for himself or herself, or for another person); **NOTE 2:** These "subject-specific" reference intervals are not addressed in this guideline.

reference sample group – an adequate number of persons selected to represent the reference population.

reference value – the value (test result) obtained by the observation or measurement of a particular type of quantity on a reference individual; **NOTE:** Reference values are obtained from a reference sample group.

The following scheme demonstrates the relationship between the terms defined.



trueness (of measurement) – closeness of agreement between the average value obtained from a large series of results of measurements and a true value (ISO 17511).[9]

4.3 Abbreviations/Acronyms

ALT	alanine aminotransferase
AST	aspartate aminotransferase
CI	confidence intervals
EPTRV	Expert Panel on Theory of Reference Values

ICSH	International Council for Standardization in Haematology
IFCC	International Federation of Clinical Chemistry and Laboratory Medicine
IQR	interquartile range
IUPAC	International Union of Pure and Applied Chemistry
JCTLM	Joint Committee for Traceability in Laboratory Medicine
K-S	Kolmogorov Smirnov
MAD	median absolute deviation about the median
M-W U	Mann-Whitney U
NTproBNP	N-terminal prohormone brain natriuretic peptide
QC	quality control
SI Units	Système International d'Unités
S-T	Siegel-Tukey
WHO	World Health Organization

4.4 Clarifications

Reference values may be associated with good health or with other physiological or pathological conditions, and they may be used for different reasons. In all cases, the reference values allow one to relate or compare observed data to reference data from a defined population of subjects. This comparison then becomes part of the decision-making process regarding the meaning of the observed value and the condition of the subject being tested.

The reference values are all values obtained by observation or measurement on reference individuals in the reference sample group. The reference interval is usually the central interval of values bounded by the reference limit values at certain designated percentiles. That is, the reference interval refers to that interval set of values observed in the reference sample group or predicted for the reference population, defined by a specific percentage (for example, the central 95%).

5 Use of Système International d'Unités (SI Units)

Although CLSI documents generally use units that are fully acceptable within the Système International d'Unités (SI), these do not always coincide with the units recommended by the International Union of Pure and Applied Chemistry (IUPAC) and by IFCC for reporting results of clinical laboratory measurements. CLSI documents also include the IUPAC/IFCC-recommended units of volume (L) and substance (molecular) concentration (mol/L) in parentheses, where appropriate.

6 Protocol Outline for Obtaining Reference Values and Establishing Reference Intervals

6.1 New Analyte or Analytical Method

The production of health-associated reference values and the subsequent estimation of the reference interval for a given analyte must be carried out in accordance with a well-defined protocol. This involves following a sequence of operations as outlined here. This outline should be applied when establishing reference values for a new analyte, for a different group of individuals, or for a new analytical method with improved analytical sensitivity and specificity for a previously measured analyte:

(1) Establish a list of analytical interferences and sources of biological variability from medical and scientific literature (in the case of a totally new analyte, the literature may not be helpful, which necessitates a new laboratory investigation of these matters).

(2) Establish selection (or exclusion) and partition criteria and an appropriate questionnaire designed to reveal these criteria in the potential reference individuals.

(3) Execute an appropriate written consent form for participation in the reference interval study and have the reference individual complete the questionnaire.

Even though performing a reference interval study is not—strictly speaking—research, questionnaires, consent forms, and even the nature of the exercise may need to be reviewed by the institution's Internal Review Board or Human Subjects Committee. Laboratories are urged to familiarize themselves with their local policies.

(4) Categorize the potential reference individuals based on the questionnaire findings and results of other appropriate health assessments.

(5) Exclude individuals from the reference sample group based on the exclusion criteria or assessments indicating a lack of good health.

(6) Decide on an appropriate number of reference individuals in consideration of desired confidence limits.

(7) Prepare, properly and consistently, the selected persons for specimen collection for the measurement of a given analyte consistent with the routine practice for patients.

(8) Collect and handle the biological specimens properly and in a manner consistent with the routine practice for patient specimens.

(9) Collect the reference values by analyzing the specimens according to the respective analytical methodology under well-defined conditions and consistent with the routine practice for patient specimens.

(10) Inspect the reference value data and prepare a histogram to evaluate the distribution of data.

(11) Identify possible data errors and/or outliers.

(12) Analyze the reference values, ie, select a method of estimation and estimate reference limits and the reference interval (include partitioning into subclasses for separate reference intervals, if appropriate).

(13) Document all of the previously mentioned steps and procedures.

The previous sequence of operations is consistent with the *a priori* approach (see Section 7.4) of selecting reference individuals and determining reference values. As a practical matter, when examining groups of potential reference individuals that are expected to be healthy, the questionnaire completion and specimen collection are often executed at the same time. The analytical measurement should be cancelled in the case of a discovered exclusion.

In some cases (see Section 7.4), the *a posteriori* method may be useful or even necessary. This approach uses measured values from a large collection of data already obtained on medically examined or otherwise grouped persons. For the *a posteriori* method, the same considerations for including certain persons and their respective measured values as reference values must be made but only after the measurements are taken rather than before.

6.2 Multicenter Reference Interval Studies

Several variables influence a reference interval, among them the criteria used to define the reference individuals, preanalytical conditions, and statistical treatment of the data. However, two variables are particularly important: the analytical method used, and the population from which the reference individuals are taken. The working group notes the recommendation that each laboratory determine its own reference intervals derives principally from these two variables.

With ongoing international efforts to standardize methods, the effects of site-to-site differences in the implementation of many analytical methods should be minimized, if not largely eliminated. Through the efforts coordinated by the Joint Committee for Traceability in Laboratory Medicine (JCTLM), analytes for which an official "reference measurement system" exists should reach a reasonable level of comparability independent of analyzer, reagent, and even analytical principle (provided the different principles have similar specificity). (A database of such analytes is available at http://www.bipm.org/jctlm/.)

Once comparability in methods is achieved, the only remaining reason for each laboratory to determine its own reference intervals is the assumption that there are differences in the reference population. There are, of course, some examples of such differences, including the effect of race on serum creatinine[13] and the effect of region on some specific proteins in Asian populations.[14] For many, if not most, analytes, however, there are few data documenting such differences between populations.[14]

Thus, it may be possible to produce a common "set" of reference intervals through a multicenter effort. In order to perform a multicenter reference interval study, the following criteria need to be satisfied:

- *A priori* selection of reference subjects according to the points outlined in the previous section. The number of participating centers, and the number of enrolled individuals, should be commensurate to the number of subjects required to allow partitioning by age, sex, race, etc.

- Clear definition of the preanalytical phases (see Section 8).

- Demonstration of traceability of results and interlaboratory standardization, ideally by inclusion of two (or more) fully commutable reference materials (frozen pools) with target values assigned by a reference method. This part is extremely critical, because it guarantees the traceability to higher order references and thus a worldwide applicability.

- Well-defined QC program with clear criteria, defined *a priori,* for acceptance or rejection of each laboratory's analytical data.

Differences in populations could be discerned from these data (see Section 9.3). For populations without such differences, the data could be pooled, allowing large numbers of observations to be analyzed. For populations with such differences, the differences would thus be well documented.

Once the common set of reference intervals is ***established***, each individual laboratory then has only to ***validate*** these reference intervals in its own environment (see Section 11).

6.3 Previously Measured Analyte

When a valid reference value study exists, it may be preferable to transfer a reference interval without having to perform a new, full-scale study. Transference can only be deemed acceptable if the test subject population, and the entire methodology, from preparation of the test individual to the analytical measurement, are the same or appropriately comparable. See Sections 10 and 11 for details on transference and validation of existing reference intervals.

7 Selection of Reference Individuals

7.1 Introduction

This section provides guidelines and suggestions for obtaining a *reference sample group* of *reference individuals* from a *reference population.*[2,3] Section 4.2 of this document gives definitions of the above italicized terms. In this section, the concepts of exclusion and partitioning are explored, a sample questionnaire to facilitate the selection and categorization processes is presented, and different sampling techniques are described.

The major intent of this document is to present procedures for determining "health-associated" reference values. Health is a relative condition lacking a universal definition. Defining what is considered healthy becomes the initial problem in any study, and establishing the criteria used to exclude the nonhealthy from the reference sample is the first step in selecting reference individuals. Each institution or investigator may have different criteria for health; these criteria should be defined before proceeding. The designation of good health for a candidate reference individual may involve a variety of examinations, such as a history and physical and/or certain clinical laboratory tests. The criteria used for any reference value study should be described and documented so others can evaluate the health status of that reference sample group. At a minimum, a questionnaire (see Section 7.3) should be used to evaluate the health of each reference individual.

7.2 Exclusion and Partitioning

Exclusion criteria are details about the candidate reference individual that, if present, serve to keep that person from being included in the reference sample. Examples of some potential exclusion criteria are found in Table 1. Certain items in Table 1 may need to be controlled when selecting persons for a reference sample for health-related reference intervals. Table 1 is not exhaustive but should serve to stimulate thinking about criteria. Not all reference value studies have the same exclusion criteria.

Table 1. Examples of Possible Exclusion Criteria

Alcohol consumption	Illness, recent
Blood donor	Lactation
Blood pressure, abnormal	Obesity
Drug abuse	Occupation
Drugs, prescription	Oral contraceptives
Drugs, over the counter	Pregnancy
Environment	Surgery, recent
Fasting or nonfasting	Tobacco use
Genetic factors	Transfusion, recent
Hospitalization, current/recent	Vitamin abuse

Partitioning criteria are characteristics of the selected reference individual that divide the reference sample into significant subclasses. Two of the most common partitioning criteria are age and sex. Table 2 lists others. Again, this is not intended to be an exhaustive list, but rather it should stimulate thinking about the partitions appropriate for the reference interval study being designed.

Table 2. Examples of Possible Partitioning Factors

Age	Geographic location
Blood group	Posture when sampled
Circadian variation	Race
Diet	Sex
Ethnic background	Stage of menstrual cycle
Exercise	Stage of pregnancy
Fasting or nonfasting	Tobacco use

What may be considered an exclusion criterion in one study could be used to partition in another. An example of this might be pregnancy. A laboratory serving a general population may choose to exclude pregnant women from their reference sample; however, a laboratory that supports an obstetrics group practice may choose to partition its samples from reference pregnant women by trimesters.

Well-designed questionnaires provide an excellent way to implement the exclusion and partitioning criteria. These forms should be simple and nonintimidating. Questions should most often require yes or no answers and simple, explanatory responses. The questionnaire may be used in conjunction with some simple measurements, such as blood pressure, height, and weight, and also with an interview where it is appropriate to ask interviewees if they consider themselves to be in good health. Common sense should apply when evaluating the responses. A sample questionnaire is included as part of Section 7.3.

7.3 Sample Questionnaire

The questionnaire is presented in this document as an example (see Figure 1). To protect the reference individuals, it is important to maintain the questionnaire information and the testing results in a confidential manner. Several design changes might be considered.

Name, address, and phone number are included to facilitate contacting the reference individual in case the analysis uncovers some potential abnormalities. In such cases, the working group believes there is an obligation to notify the person or his or her physicians. The laboratory should have a mechanism in place for medical review and confidential notification.

In some situations, anonymous questionnaires may be a better vehicle for obtaining the required information. In these instances, a numbering system could be used. In this case, it is the responsibility of the reference individual to contact the laboratory to determine if the testing showed any problems that require follow-up. In this sense, the anonymous questionnaire approach is more problematic.

Another possible variation, especially in the case of an *a priori* study (see Section 7.4.1), is to group the questions by exclusion and partitioning. Questions that are designed to uncover information about disease states known to affect the tests under investigation should be included.

It is appropriate that the laboratory obtain written informed consent from each reference individual. The consent form should state clearly that laboratory personnel are allowed to obtain specimens, and to use the associated laboratory values and questionnaire information for the determination of reference intervals. Usually, the informed consent accompanies the questionnaire. Even though performing a reference interval study is not—strictly speaking—research, questionnaires, consent forms, and even the nature of the exercise may need to be reviewed by the institution's Internal Review Board or Human Subjects Committee. Laboratories are urged to familiarize themselves with their local policies.

ALL INFORMATION IS STRICTLY CONFIDENTIAL AND IS FOR USE WHEN DIAGNOSING ILLNESS AMONG MEMBERS OF YOUR COMMUNITY.

SUBJECT ID #_____SAMPLE ID #

NAME:_____PHONE_____
LAST FIRST MIDDLE

ADDRESS:_____

AGE:_____(YRS) SEX: (M) (F) RACE: _____

HEIGHT:_____ FT _____ IN WEIGHT:_____ LBS

OCCUPATION:_____

PHYSICIAN NAME:_____

1. DO YOU CONSIDER YOURSELF TO BE HEALTHY? (Y) (N)

2. DO YOU EXERCISE REGULARLY? (Y) (N)
IF YES, HOW OFTEN? (HRS PER WK)_____
AND DEGREE OF ACTIVITY? (LIGHT) 1 2 3 4 5 6 7 8 9 10 (VIGOROUS)

HAVE YOU BEEN SICK RECENTLY? (Y) (N)
IF YES, WHEN?_____DESCRIBE ILLNESS: _____

3. ARE YOU TAKING ANY PRESCRIBED MEDICATION? (Y) (N)
IF YES, WHAT?_____

4. DO YOU HAVE HIGH BLOOD PRESSURE? (Y) (N)

5. DO YOU TAKE VITAMIN SUPPLEMENTS OR HERBAL REMEDIES? (Y) (N)
IF YES, WHAT?_____

6. ARE YOU EXPOSED TO ANY HAZARDOUS CHEMICALS IN YOUR JOB? (Y) (N)
IF YES, WHAT?_____

7. DO YOU USE TOBACCO? (Y) (N)

IF YES, WHAT FORM?_____HOW OFTEN?_____

8. DO YOU EAT A SPECIAL DIET? (Y) (N)

IF YES, PLEASE DESCRIBE_____

Figure 1. Sample Questionnaire

9. DO YOU DRINK ALCOHOLIC BEVERAGES? (Y) (N)

IF YES, WHAT FORM?_____HOW OFTEN?_____

10. ARE YOU CURRENTLY UNDER A DOCTOR'S CARE? (Y) (N)

IF YES, WHY?_____

11. HAVE YOU BEEN HOSPITALIZED RECENTLY? (Y) (N)

IF YES, WHY?_____WHEN?_____

12. ARE THERE ANY INHERITED HEALTH DISORDERS IN YOUR FAMILY? (Y) (N)

IF YES, DESCRIBE:_____

13. HAVE YOU TAKEN ASPIRIN OR ANY PAIN RELIEVERS RECENTLY? (Y) (N)

IF YES, WHAT?_____WHEN?_____

14. HAVE YOU TAKEN ANY COLD OR ALLERGY MEDICINE RECENTLY? (Y) (N)

IF YES, WHAT?_____WHEN?_____

15. HAVE YOU TAKEN ANY ANTACIDS OR STOMACH MEDICINE RECENTLY? (Y) (N)

IF YES, WHAT?_____WHEN?_____

16. ARE YOU TAKING DIET PILLS? (Y) (N)

FOR WOMEN:

1. ARE YOU STILL MENSTRUATING? (Y) (N)

IF YES, WHEN WAS YOUR LAST PERIOD?_____

IF NO, ARE YOU ON HORMONE REPLACEMENT THERAPY? (Y) (N)

2. ARE YOU BREAST-FEEDING? (Y) (N)
3. ARE YOU PREGNANT? (Y) (N)

IF YES, WHAT IS YOUR DUE DATE?_____

4. ARE YOU USING ORAL OR IMPLANT CONTRACEPTIVES? (Y) (N)

Figure 1. Sample Questionnaire (Continued)

7.4 Selection of Reference Individuals

Reference individuals for the determination of a health-associated reference interval do not necessarily have to be young adults; they may more closely resemble the patient population undergoing medical evaluation. In fact, the working group rejects, in general, the concept of an unequivocal "gold standard" of young, healthy adults and suggests that age-related reference intervals, in many instances, may be more clinically appropriate. However, the working group also acknowledges that some age-related changes in laboratory values may not represent good health (eg, increases in alkaline phosphatase in the geriatric patient).

The working group strongly endorses the use of **direct sampling techniques**, in which reference individuals are selected from a reference population using specific, well-defined criteria. When these criteria are applied before samples are collected and analyzed, it is referred to as ***a priori.***

If these same criteria are applied following sample collection, it is referred to as ***a posteriori.***

However, the working group also acknowledges that, in some circumstances (eg, pediatrics), it may be particularly difficult to use direct sampling techniques. In these cases, some investigators have advocated the use of **indirect techniques,** in which individuals are not considered but certain statistical methods are applied to values in a database. The working group does **not** endorse these methods as primary approaches for establishing reference intervals.

7.4.1 Direct Sampling Techniques

A priori sampling is a method that requires well-defined exclusion and partitioning criteria before the selection of the reference individuals. This is a method best applied to well-studied, established laboratory procedures. With established methods, a thorough search of the literature should identify known sources of biological variation. The information from the literature is then translated into a list of exclusion and partitioning criteria appropriate for the study under development. After these criteria are established, a questionnaire is typically developed to use in conjunction with an interview to exclude certain persons from the sampling process and partition selected persons into their subclasses. This entire process takes place before any blood samples are collected. The number of reference individuals selected for analysis must be an adequate number to be statistically valid (see Section 9.1).

In *a posteriori* sampling, the process of exclusion and partitioning also takes place but in a different order (ie, after sampling and analyte testing rather than before). The *a posteriori* approach may be especially appropriate for laboratory procedures that are new or poorly studied, and for which the literature contains little information. Because the factors defining a subclass may not be known initially, the questionnaire for this approach may need to be more thorough than the one designed for the *a priori* sampling process.

7.4.2 Indirect Sampling Technique

In indirect sampling techniques, laboratory values from a database established for other purposes (eg, a standard laboratory information system) are used for estimating reference intervals. These techniques are used when it is deemed too difficult to collect samples from healthy subjects (eg, pediatrics). Although this approach is relatively simple and relatively inexpensive, one must take extra precautions not to include large numbers of values from unhealthy individuals who may be present in the database.

The indirect sampling techniques are based on the assumption, confirmed by observation, that most results, even on hospital and clinic patients, appear "normal." Several methods are used to exclude values from unhealthy individuals, and statistical approaches are available to extract reference values from hospital data.[15-20] In many studies, data from all hospital patients (or all outpatients) are used to estimate reference intervals,[15,20] but the techniques are perhaps more appropriately employed using data from individuals who are relatively healthy:

- blood donors;
- individuals undergoing routine physical examinations for periodic health screening;
- individuals undergoing lead screening;
- patients undergoing minor surgical procedures; and
- individuals undergoing genetic screening (eg, unaffected parents and siblings of a patient with cystic fibrosis).

Once the data are extracted from the study population by applying exclusion and partitioning criteria, statistical methods described in other sections (see Section 9) can be used to estimate reference intervals.

No matter how they are calculated, reference intervals generated with indirect techniques should be considered rough estimates at best, as the underlying assumption that most of the data come from reference individuals may not be correct.

Whenever possible, the working group recommends using direct methods over indirect methods for establishing and verifying reference intervals.

8 Preanalytical and Analytical Considerations

Analytical results from reference populations must reflect all of the preanalytical and analytical variables that can influence test results. Therefore, all preanalytical factors, including subject preparation, sample collection and processing, the analytical method, and instrumentation, must be carefully defined and used for testing both reference individuals and the patient population.[3,21]

Control of clinically meaningful, preanalytical factors is essential to minimize the effect on clinical decision making. Therefore, it may be necessary, for certain analytes, to define conditions for establishing reference intervals in different subclasses (eg, hospitalized recumbent patients vs ambulatory outpatients or specimens drawn in the morning vs specimens drawn in the afternoon). Many of these preanalytical situations constitute partitioning factors, such as those described in Section 7.2, and they may require separate reference intervals. In some cases, the laboratory and physician have some control over the preanalytical variables, which obviates the need to separate the reference intervals.

In general, preanalytical considerations involve two areas, namely, biological and methodological factors.[22] The biological factors include those that are of metabolic and hemodynamic origin. Procedures resulting in potential for cell damage (from physical training to venipuncture) should be considered. Subjects using pharmacologic agents causing enzyme induction should have already been excluded. The preanalytical methodological factors involve specimen collection and handling, including consideration of collection techniques, additives, and the order of filling the tubes (for blood samples). Refer to the IFCC checklist [3] and Table 3 for helpful guidelines for evaluating preanalytical issues.

Measurement of the same analyte by more than one method, instrument, or system requires appropriate examinations to verify that the various methods, instruments, or systems generate comparable results. If the alternate methods or systems do not give comparable results (see Section 10 and CLSI/NCCLS document EP09[8]), then separate reference intervals may need to be established, particularly if the differences in the numerical results are clinically significant.

8.1 Subject Preparation

As described in Section 7, the selection of reference individuals must appropriately address many issues. Inadequate subject preparation or deviations from the defined criteria may give rise to results that are inaccurate or that skew data. The criteria set are dictated by the effect of biological variation on the analyte(s) of interest. Table 3 summarizes some of the factors to consider regarding subject preparation.[3,22]

Food ingestion before blood sampling affects many laboratory results, either directly (changes in analyte concentration) or indirectly (lipid interference). Conversely, prolonged fasting induces other changes. Many analytes also are affected by common agents such as caffeine, ethanol, tobacco, and vitamin C. Therefore, use of these agents, or any others, must be evaluated as part of the patient/subject preparation scheme.[3,22]

Exercise and postural position during the phlebotomy procedure can change a laboratory result. The impact of postural changes is important when comparing inpatient and outpatient results and, as stated earlier, frequently necessitate the establishment of separate reference intervals for some analytes. Other factors to consider include ethnic background, seasonal changes, and circadian rhythms, all of which may affect analyte concentration.[21-23]

Many of these issues are eliminated by the appropriate exclusion criteria. (Solberg and PetitClerc[3] provide specific details about each category.)

Table 3. Preanalytical Factors for Consideration

Subject Preparation	Specimen Collection	Specimen Handling
• Prior diet • Fasting vs nonfasting • Abstinence from pharmacologic agents • Drug regimen • Sampling time in relation to biological rhythms • Physical activity • Rest period before collection • Stress	• Environmental conditions during collection • Time • Body posture • Specimen type • Collection site • Site preparation • Blood flow • Equipment • Technique • Tourniquet time	• Transport • Clotting • Separation of serum/plasma • Storage • Preparation for analysis

8.2 Specimen Type, Collection, Handling, and Storage

The laboratory should have a manual outlining the collection, handling, and storage of specimens so appropriate applications of reference intervals can be made by the physician when interpreting patient results. Care should be taken to specify the appropriate blood collection tubes for serum, plasma, or whole blood samples (see CLSI documents H03, H04, and H21, and CLSI/NCCLS document H11[23-26]).

Consideration should be given to whether the specimen should be maintained under anaerobic conditions (eg, for ionized calcium measurements) or collected and shipped on ice (eg, for ammonia and lactate). Knowledge of the types of evacuated tubes or syringes used to collect fluids is important. Serum or plasma separator tubes or siliconized syringes can interfere with certain tests, which could cause erroneous results. Specimen integrity must also be considered. Fluids should be clear, that is, free of red cells and other debris. The laboratorian should use discretion on some issues, and he or she may refer to the literature for information when questions arise about potential effects of deviation from the standardized protocol.

8.2.1 Blood

If blood is the specimen of choice, it is necessary to define whether the sample should be arterial, venous, or capillary; whether the specimen should be anticoagulated; and, if anticoagulated, which anticoagulant is acceptable. The conditions for standardized specimen collection by venipuncture and skin puncture are described elsewhere (see CLSI documents H03 and H04[23,24]).

8.2.2 Other Body Fluids

Specimen procurement of other body fluids, although generally not under control of the laboratory, still requires definition of specific guidelines for collection, processing, and handling. Such fluids include urine (see CLSI/NCCLS document GP16 and CLSI document C49[27,28]); cerebrospinal, pleural,

pericardial, peritoneal, synovial, and amniotic fluids; and saliva. In some instances, the drawing of a concomitant blood sample can be necessary, but in others, timed collections can be appropriate. As in the case of blood, knowledge of the use of such substances as preservatives and anticoagulants is critical. In the case of 24-hour urine collections, it is highly desirable to "validate" the completeness of the collection by determination of the total creatinine excreted.

8.2.3 Temperature

The collection and handling of some specimens may require procurement at a specific temperature (eg, 37 °C, room temperature, or iced). In addition, preservation of some specimens (analytes) requires storage at a particular temperature or freezing, possibly at a specified temperature (−20 °C vs −70 °C). It is essential to establish any special conditions and strictly adhere to them. In general, specimens should be processed promptly after collection. Processing frequently entails removal of serum or plasma from the clot or red cells as quickly as possible and at a specified temperature[22] (see CLSI/NCCLS document H18[29]).

8.3 Analytical Method Characteristics

Besides intra- and interindividual variability, the reference intervals also include the analytical variability of the method used for measurement. Thus, the validity of information provided by the laboratory is critical. The methods used must be described in detail, reporting between-run analytical imprecision, limit of detection, linearity, recovery, and interference characteristics, and especially its trueness and the demonstration of traceability of the results provided to higher order methods or materials, when they exist according to ISO 17511[9] and CLSI document X05.[7]

Other factors that affect analytical performance require consideration. These include equipment/ instrumentation, reagents (including water), calibration standards, and calculation methods. The establishment of reference intervals must also include lot-to-lot and technologist variability, as well as instrument-to-instrument variability if duplicates of the same analyzer are used. Knowledge of all the above factors defines the analytical system to be checked.

The reliability of the data generated is critical, because both the precision and trueness of the method determine its diagnostic utility. Therefore, also included during the determination of reference intervals is the routine use of QC materials in the same format as for patient testing, which not only monitors the analytical protocol used during the process, but also ensures equivalence of results over the long term. (Refer to CLSI document C24.[30]) Ideally, data are gathered by analyzing specimens over several days, resulting in values that represent average run-to-run variation. In addition, an assessment of the interference from naturally occurring constituents is essential.[31]

9 Analysis of Reference Values

The reference interval is defined here as the interval between and including two numbers, an upper and lower reference limit, which are estimated to enclose a specified percentage (usually 95%) of the values for a population from which the reference subjects are drawn. For most analytes, the lower and upper reference limits are estimated as the 2.5th and 97.5th percentiles of the distribution of test results for the reference population, respectively. In some cases, only one reference limit is of medical importance, usually an upper limit, say the 97.5th percentile.

The confidence intervals for the estimates of the limits of the reference interval can be constructed assuming random sampling of the reference population. The width of each confidence interval depends on the number of reference subjects, as well as the distribution of the observed reference values.

In the previous version of this document, two general statistical methods for determining such limits were described: the nonparametric and the parametric procedures. Full details of these procedures are given in

Part 5 of the published documents of the EPTRV prepared by Solberg.[5] The nonparametric method of estimation makes no specific assumption about the probability distribution of the observed reference values. The parametric method, as applied in practice, assumes that the observed values, or some mathematical transformation of those values, follow a Gaussian (ie, "normal") probability distribution. Because the reference values of many analytes do not follow the Gaussian distribution, use of the parametric method requires that they be transformed to "normalize" them. This requires selecting the most suitable transformation (eg, logarithmic, power, or some other function) and then testing whether the transformed reference values conform to a Gaussian distribution. This involves some moderately complex statistical theory and corresponding computer programs. An excellent, detailed discussion of these matters is contained in Appendixes B and C of the EPTRV publication.[5]

The simple nonparametric method remains the recommended procedure for establishing reference intervals if a laboratory has limited access to statistical and computational support. The working group emphasizes that the most important considerations in developing reliable reference intervals are selecting appropriate reference subjects, testing an adequate number of subjects, and avoiding preanalytical errors, not the statistical method used to estimate the reference intervals from the observed data. If sample size constraints of the simple nonparametric method prevent a laboratory from establishing reference intervals, and the laboratory has access to personnel that can interpret and implement more complex procedures, the working group recommends use of either bootstrap-based procedures,[32-35] traditional parametric methods,[5,32,33] or the robust methods described in Horn and Pesce, 2005.[36]

Among these three latter methods, the working group would like to call attention to the "robust method," because it may offer the best way of dealing with limited numbers of observations. The robust method can be thought of as a compromise between the parametric and nonparametric methods. It has the appeal of the parametric method in that it does not require as many observations as the nonparametric procedure, and yet it does not require that the underlying population of analytical values follow a Gaussian distribution. This method has the same form as the parametric except, instead of the mean and standard deviation of the sample, robust measures of location and spread are used. The robust method has been used in a variety of situations where the available sample size is less than 120, but where the underlying population is not assumed to follow a Gaussian distribution. Details on the computations involved may be found in Horn and Pesce, 2005.[36]

Examples of the nonparametric method of estimating reference intervals for two analytes, serum calcium and alanine aminotransferase (ALT), are described in Section 9.4.1. In addition, a demonstration of the use of the robust method on subsets of those data is provided in Section 9.4.2 and a detailed example is given in Appendix B.

9.1 Minimum Number of Reference Values

Using the nonparametric method, it is impossible to distinguish between two percentiles of a distribution that are P% apart unless at least $n = (100/P) - 1$ observations are obtained. The reason for this is that the nonparametric method is based solely on the ranks of the observations (in order of magnitude) and ignores their measured values. For example, if a sample of nine observations is taken at random from some population, only nine estimates of percentiles can be obtained from the nine rankings when these are ranked in order of magnitude. The smallest observation is the nonparametric estimate of the 10th percentile of the population; the largest observation is the nonparametric estimate of the 90th percentile of the population. Thus, as the formula states, a sample of nine observations [$9 = (100/P) - 1$, where $P = 10.0$] represents the minimum sample size necessary to obtain distinct nonparametric estimates of the ordered population deciles, which are, by definition, percentiles of the population exactly 10% apart from each other.

Similarly, to estimate the 2.5th percentile distinct from the 5th percentile, or the 95th percentile distinct from the 97.5th (ie, $P = 2.5$), a minimum of 39 measurements is required [$39 = (100/2.5) - 1$]. The

smallest observation in the sample is the nonparametric estimate of the 2.5th percentile of the population, while the largest observation estimates the 97.5th percentile.

It certainly is undesirable, however, to rely entirely on the extremes of a set of observed values in order to derive a nonparametric 95% reference interval. These might be aberrant or otherwise nonrepresentative of the true percentile values of the population. Reed et al[37] suggest that a minimum of 120 observations be secured, one from each reference subject. This has the advantage of also allowing 90% confidence limits to be computed nonparametrically for each reference limit (see Section 9.5). To estimate the reference limits for these same percentiles with 95% confidence, a minimum of 146 reference values are needed; for 99% confidence, a minimum of 210 reference values are needed. Linnet[38] proposes that up to 700 should be obtained for highly skewed distributions of results. However, as a standard for general practice, the working group supports the recommended minimum of 120 reference subjects.

This number assumes that no observations are deleted from the reference set (see Section 9.2). If aberrant or outlying observations are deleted, then additional subjects should be selected until at least 120 acceptable reference values are obtained for each determination of a reference interval. Moreover, if separate intervals are needed for different subclasses (by sex or age, for example), each such interval should be based on the recommended number (at least 120) of reference observations.

In the case of subclass reference values for certain populations, such as newborn, pediatric, and geriatric patients, it may be difficult, if not impossible, to obtain appropriate reference subjects in sufficient numbers. Whatever number of values is obtained, the data should still be analyzed by the nonparametric method and reported by percentiles appropriate to the number of values obtained. As an alternative, the robust method may be used in these situations.

With respect to the robust method, there is no specific minimum number of required observations. It is, of course, desirable that as many good observations as possible be used in the calculation of the reference interval. The drawback to using fewer observations is the statistical uncertainty for the smaller sample sizes, translating into wider confidence intervals for the limits of the reference intervals (see Section 9.5). This type of uncertainty decreases as the number of subjects increases.

In summary, the method for establishing reference intervals, the shape of the reference population distribution (approximately Gaussian vs highly skewed), and the tolerable uncertainty in the reference limits all influence the minimum required number of reference values. A general criterion for determining sample size is that the width of the 90% confidence interval (CI) for a reference limit should be acceptably small relative to the width of the 95% reference interval itself (Harris and Boyd)[33]; these authors recommend that the width of the 90% CI be less than 0.2 times the width of the reference interval. If CIs are unacceptably wide, more reference values are needed. While the selection of estimation method can have some impact on the width of CIs for a given reference distribution, the factor that influences the width of the CIs most importantly is the number of available reference values.

9.2 Treatment of Outlying Observations

An important implicit assumption in the estimation of reference limits is that the set of measured reference values represents a “homogeneous” collection of observations. This means all values come from the same underlying population of test results characterized by a probability distribution (even though, under the nonparametric method, the form of this distribution is not specified).

It may be that this condition is satisfied by most of the reference values, but some arise from a different population of test results. When such values lie in the midst of the others, they are practically impossible to identify, unless the person performing the biochemical analysis happens to know that these observations represent atypical analytical conditions or are the result of some arithmetic or procedural mistake.

Often, however, such reference values arising from a different population of test results lie outside the range of the bulk of the reference values, and they are easily identified as "outliers" requiring special attention.

Unless outliers are known to be aberrant observations (eg, due to a mistake in the analysis or to a lapse in the preanalytical controls applied to the remaining subjects), the emphasis should be on retaining rather than deleting them. Nonparametrically estimated reference limits based on at least 120 observations are not changed at all, if an extreme value is deleted.

Thus, among the first steps in the analysis of data collected in connection with a reference interval study is a **visual** examination of the frequency distribution.

Many statistical techniques are available for detection of outlying observations (see Barnett and Lewis[39]). The majority of these techniques rest on the assumption that the observed reference values are Gaussian-distributed. Moreover, when any test for outliers is performed on extreme values individually, there is always a possibility that less extreme outliers may be masked.

A test proposed by Dixon[40] has become fairly well known in reference value estimation, namely, the ratio *D/R*, where *D* is the absolute difference between an extreme observation (large or small) and the next largest (or smallest) observation, and *R* is the range of all observations, including extremes. Reed et al[37] suggest the value one-third as a cutoff value; ie, if the difference *D* is equal to or greater than one-third of the range *R*, the extreme observation is deleted. For sample sizes as large as 120, this criterion is rather conservative (as Reed et al point out). That is, less than 1% of the time it would eliminate as an outlier an observation that belongs to the same underlying (normal) distribution as the rest of the observations. However, in the absence of evidence that an outlier is indeed an aberrant observation, and given that the underlying distribution often is not exactly Gaussian in form, the one-third rule for the ratio *D/R* seems appropriate, especially when reference intervals are determined by the nonparametric method. Therefore, the working group supports the use of this test and cutoff value in examining a set of observed reference values for statistically significant outliers.

When two or three outliers are present on the same side of the distribution (ie, all extremely large or extremely small), the one-third rule (or any similar *D/R* rule) may fail to label the most extreme outlier as statistically significant, and thereby mask the presence of the other outliers that are just slightly less extreme. In such a case, the one-third rule should be applied to the least extreme outlier as if it is the only outlier. If the rule leads to rejection of this outlier, then the more extreme observations should naturally be rejected, as well. If the rule does not reject the least extreme value, then one should either accept all the extreme values or, alternatively, apply a test that considers all the outliers together. Such a test is called a *block* procedure; examples are given in Barnett and Lewis.[39]

Another method of outlier detection, proposed by Tukey (1977),[41] uses only the middle 50% of the sample, thus reducing, or even eliminating, the possible masking effect of multiple outliers on one side of the distribution. This method involves the computation of the lower and upper quartiles (ie, the 25th and 75th percentiles) of the sample dataset; call these statistics Q_1 and Q_3. Next, the interquartile range, Q_3-Q_1, or IQR, is computed. Lastly, the lower and upper "boundaries" are computed as follows: the lower boundary = $Q_1 - 1.5 \times IQR$ and the upper boundary = $Q_3 + 1.5 \times IQR$. Any data point outside the boundaries (ie, either less than the lower boundary, or greater than the upper boundary) is considered an outlier, and is omitted from subsequent reference interval estimation. In theory, this method on average excludes roughly 0.7% of the data belonging to the Gaussian distribution. When reference values are not Gaussian distributed, they need to be transformed. A reasonable family of transformations is the power family due to Box and Cox (1964):[42]

$$y = \begin{cases} \dfrac{x^{\lambda} - 1}{\lambda} & \lambda \neq 0 \\ \ln(x + c) & \lambda = 0 \end{cases}$$

where y is the transformed value, x is the original value, and ln is the natural logarithm. The parameters λ and *c* are estimated using maximum likelihood techniques. Details on the use of the Box-Cox transformation in conjunction with the Tukey outlier detection technique are found in Horn and Pesce (2005).[36]

When any outlier is rejected, it is appropriate to test the remaining data for an additional outlier(s).

These same techniques should be applied regardless of which method is ultimately used to generate reference intervals.

9.3 Partitioning of Reference Values

The possibility that separate intervals are desired for defined subclasses of subjects should be considered before the actual process of securing and analyzing subject specimens. Separate reference intervals for men and women or for different age groups may not be justified unless they are clinically useful and/or are well grounded physiologically. Of course, the information necessary to decide these questions may not be available for a new analyte. However, if these conditions are satisfied, then at least 120 subjects of each sex or age or other subclass should be sampled.

It is generally assumed that as long as the difference between the observed means of two subclasses is statistically significant (at the 5% or 1% probability level), then each subclass warrants its own reference interval. However, any observed difference, no matter how unimportant clinically, tests out statistically significant if the sample sizes are large enough. Sinton et al[43] have suggested that separate reference intervals not be estimated unless the difference between the subclass means is at least 25% as large as the 95% reference interval estimated from the combined (overall) sample of reference subjects.

Harris and Boyd[44] have shown, though, that smaller differences between subclass means can lead to situations in which the proportions of a subclass outside the reference limits (without partitioning) are much different from the desired 2.5% on each side. It is their contention that, when the percentage of individuals in a subclass outside one of the reference limits for the combined groups exceeds 4%, the resulting differences in sensitivity and specificity for that subclass may seriously hamper the interpretation of laboratory results as part of the diagnostic process.

Furthermore, these investigators demonstrated that these same problems can occur even when the means are identical. If the standard deviations of the subclasses are in a ratio of 1.5 or more, a larger proportion of the wider distribution (subclass) extends beyond the narrower distribution on both sides.

Thus, before the actual sampling of reference subjects is undertaken, the possibility of subclass reference intervals with respect to the analytes concerned should be considered. Pertinent physiological information on each analyte and the potential usefulness of separate subclass intervals in clinical practice should be evaluated at that time.

If such evaluation indicates that subclass distinctions may exist and may be of clinical significance, at least 120 reference subjects in each subclass should be sampled. To assist in deciding whether to partition reference intervals by subclass, Harris and Boyd[44] suggested calculating the statistical significance of the difference between subclass means by the standard normal deviate test:

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\left[\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)\right]^{1/2}} \quad (1)$$

where $\bar{x}_1$ and $\bar{x}_2$ are the observed means of the two subgroups, s_1^2 and s_2^2 are the observed variances, and n_1 and n_2 are the number of reference values in each subclass, respectively.[44] If the original data are highly skewed, and a simple transform, like the log transform, for example, produces a distribution of values much closer to Gaussian form, then it is preferable to apply the *z*-test to the transformed values.

Harris and Boyd compared the calculated statistic *z* with a "critical" value[44]:

$$z^* = 3(n_{average}/120)^{½} = 3[(n_1 + n_2)/240]^{½}. \quad (2)$$

If the calculated *z* exceeds *z**, they recommend partitioning. In addition, they recommend partitioning if the larger standard deviation, for example s_2, exceeds $1.5s_1$, or equivalently, whether $s_2/(s_2 - s_1)$ is less than 3.[44]

Despite its computational simplicity, several weaknesses in the Harris/Boyd approach are documented.[45] The approach assumes the data follow a Gaussian distribution, and it cannot account for unequal prevalence of subclasses. In addition, the estimated z-score, calculated from means and standard deviations, does not necessarily reflect the tail-behavior of the underlying distributions. For example, two distributions with identical lower reference limits and unequal standard deviations can have markedly different upper reference limits and vice versa. To overcome these limitations, an alternative method has been proposed that is based upon direct estimation of the proportions of two subclasses outside the reference limit at each end of the combined distribution.[46] Lahti has provided instructional examples comparing the various methods using data from the Nordic Reference Interval Project.[47]

When more than two subclasses are compared, the issues become even more complicated, and the aid of a statistical consultant should be sought. An example of this situation appears in Harris et al.[48] Lahti[45,46] presents evidence that neither the Harris/Boyd approach nor the alternative approach seems ideal to solve a partitioning problem involving several subclasses. Further research appears to be needed in this area.

The statistical tests and criteria recommended above may also be applied to the question of whether reference intervals determined in one laboratory should be transferred without change for use in another laboratory (see Section 10).

9.4 Examples

The histograms depicted in Figures 2 and 3 show frequencies of reference values of calcium and ALT, respectively, measured in serum samples from medical students at the University of Virginia in 1987 and 1988. The frequencies are also listed in rank order in Tables 4 and 5, with a total of 120 results for each analyte from each of two subclasses, men and women, within the age group of 20 to 30 years. The histograms for calcium show roughly Gaussian distributions; whereas those for ALT show considerable skewness to the right. The extreme value of 65 U/L of ALT (Table 6) among women does not violate the one-third rule for outliers [(65−47)/60 is less than 1/3] and should be retained. The distributions of the logarithms of the ALT values are nearly Gaussian. Results for both analytes appear generally higher in men than in women, and a statistical test of separate reference intervals by sex is of interest.

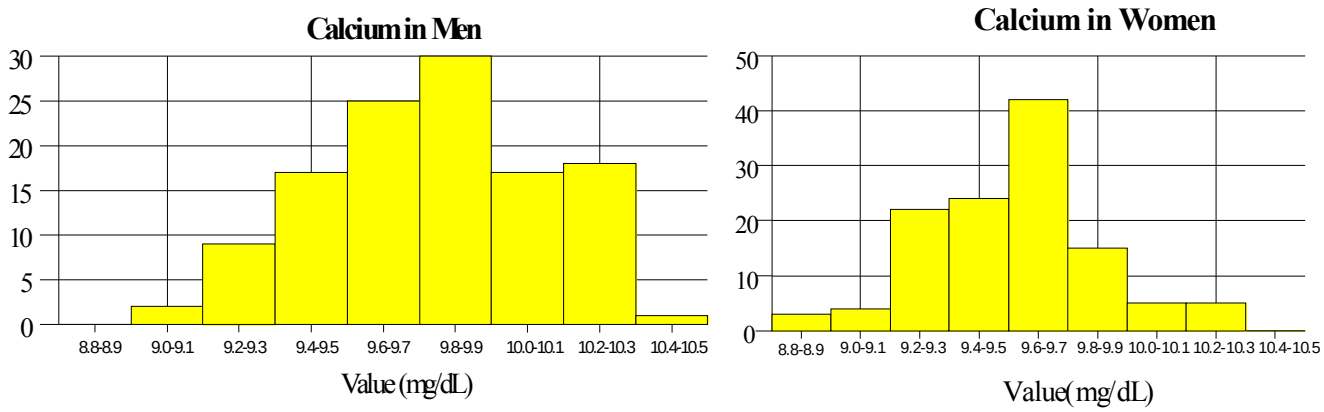


Figure 2. Calcium Histograms

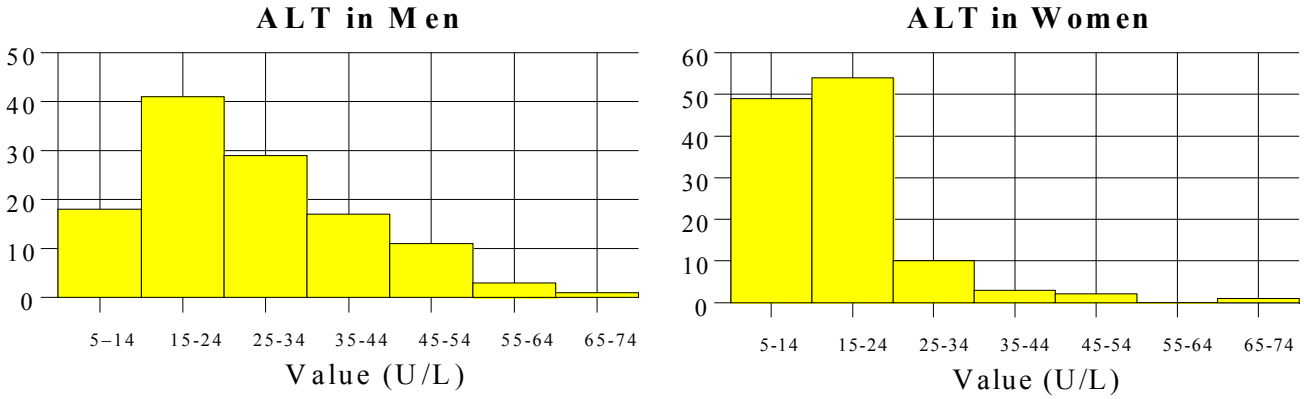


Figure 3. Alanine Aminotransferase Histograms

Table 4. Frequency Distributions of Calcium in 240 Medical Students by Sex

Value (mg/dL)*	Frequency		
	Women	Men	Combined
8.8	1	0	1
8.9	2†	0	2
9.0	1	0	1
9.1	3	2	5†
9.2	11	1†	12
9.3	11	8	19
9.4	8	6	14
9.5	16	11	27
9.6	16	12	28
9.7	26	13	39
9.8	8	16	24
9.9	7	14	21
10.0	3	7	10
10.1	2	10	12
10.2	3‡	11	14
10.3	2	7‡	9‡
10.4	0	1	1
10.5	0	0	0
10.6	0	1	1
Total	120	120	240

* mg/dL • 0.2495 = mmol/L.
† r_1 = rank value #3 (2.5th percentile), n = 120; rank value #6, n = 240.
‡ r_2 = rank value #118 (97.5th percentile), n = 120; rank value #235, n = 240.

Table 5. Frequency Distributions of ALT in 240 Medical Students by Sex

Value (U/L)	Frequency		
	Women	Men	Combined
5	1	0	1
6	3*	0	3
7	1	0	1
8	5	0	5*
9	2	1	3
10	2	2*	4
11	7	4	11
12	11	2	13
13	10	3	13
14	7	6	13
15	7	3	10
16	7	4	11
17	8	1	9
18	6	4	10
19	7	6	13
20	5	10	15
21	6	5	11
22	4	4	8
23	4	1	5
24	0	3	3
25	3	8	11
26	2	3	5
27	0	1	1
28	2	4	6
29	1	1	2
30	2	3	5
31	0	5	5
32	0	1	1
33	0	1	1
34	0	2	2
35	0	2	2
36	1	5	6
37	2	1	3
38	0	2	2
39	1	2	3
40	0	3	3
41	0	1	1
42	0	1	1
45	0	2	2
46	1†	0	1
47	1	1	2
48	0	2	2
49	0	1	1
51	0	3	3
53	0	1	1
54	0	1	1†
55	0	2†	2
62	0	1	1
65	1	0	1
69	0	1	1
Total	120	120	240

* r_1 = rank value #3 (2.5th percentile), n = 120; rank value #6, n = 240.
† r_2 = rank value #118 (97.5th percentile), n = 120; rank value #235, n = 240.

9.4.1 Nonparametric Method

Let n denote the number of observations in a set of reference data for which the 95% reference interval is computed. The observations are first ranked (ie, ordered by magnitude). Let r represent the rank of an observation (the smallest is ranked $r = 1$; the largest, $r = n$). The nonparametric method consists of computing the lower reference limit, r_1 (the 2.5th percentile), as the observation corresponding to $r_1 = 0.025\ (n + 1)$, and the upper reference limit, r_2 (the 97.5th percentile), as the observation corresponding to $r_2 = 0.975\ (n + 1)$. Since the values of r_1 and r_2 are usually not integers, the limits are generally calculated by interpolating between the data points corresponding to the ranks on either side of r_1 and r_2. However, in these examples, where $n = 120$, the values r_1 and r_2 are so close to the integers 3 and 118, respectively, that they are rounded off to these numbers:

$$r_1 = 0.025\ (121) = 3.025 \approx 3 \tag{3}$$

$$r_2 = 0.975\ (121) = 117.975 \approx 118 \tag{4}$$

For $n = 240$, the values r_1 and r_2 are 6 and 235, respectively.

Using these rank values to estimate upper and lower reference limits in the populations represented by these reference subjects, the following 95% reference intervals are obtained:

Calcium

Women:	8.9 mg/dL to 10.2 mg/dL (2.22 mmol/L to 2.54 mmol/L)
Men:	9.2 mg/dL to 10.3 mg/dL (2.30 mmol/L to 2.57 mmol/L)
Combined:	9.1 mg/dL to 10.3 mg/dL (2.27 mmol/L to 2.57 mmol/L)

ALT

Women:	6 U/L to 46 U/L
Men:	10 U/L to 55 U/L
Combined:	8 U/L to 54 U/L

To test the statistical significance of the differences between the mean values for calcium and ALT in men and women of this age group, the user needs the statistics given in Table 6.

Table 6. Means and Standard Deviations of Calcium and in ALT

Analyte	Means		Standard Deviations	
	Men (n = 120)	Women (n = 120)	Women (n = 120)	Men (n = 120)
Calcium (mg/dL)	9.80	9.57	0.31	0.29
$\log_e$ALT (ln U/L)*	3.20	2.78	0.46	0.44

* See Section 9.3 about the need for log transformation.

Inserting these statistics into equation 1 for z, the results are as follows:

$$\text{calcium}: z = \frac{|\,9.80 - 9.57\,|}{\left(\frac{(0.31)^2}{120} + \frac{(0.29)^2}{120}\right)^{1/2}} = 5.94 \tag{5}$$

$$\log_e \text{ALT}: \; z = \frac{|\,3.20 - 2.78\,|}{\left(\frac{(0.46)^2}{120} + \frac{(0.44)^2}{120}\right)^{1/2}} = 7.23 \qquad (6)$$

In both cases, the z-values exceed the critical value $z^* = 3$ for $n = 120$, indicating that separate reference intervals for men and women should be considered. However, for calcium, the clinical importance for the difference between the separate intervals is not fully understood, although on physiological grounds a significantly higher mean calcium level in young men than in young women may be expected. In a much larger sample, the difference between reference intervals for the two sexes might emerge as large enough to be more clinically useful. Therefore, for calcium, a laboratory may choose to provide a single reference range of 9.1 mg/dL to 10.3 mg/dL for both men and women in this age group. The differences may be comparable to the imprecision of the calcium analytical method, and they may be small relative to the change in calcium required for clinical significance and physician response.

For ALT, separate reference intervals by sex appear to be clinically useful for diagnostic purposes. Again, there is physiological evidence to support this conclusion.

9.4.2 Calculation of Reference Intervals on Small Sample Sizes Using the Robust Method

In Section 9.4.1, at least 120 observations were available for analysis. When fewer observations are available, use of the nonparametric method becomes problematic. The robust method, however, offers a potential alternative.

As noted in Horn and Pesce,[36] calculating the reference interval using robust methods involves an iterative process, in which the initial location (center) is estimated by the median and the initial scale (spread) by the median absolute deviation about the median (MAD). In the process, actual observations are downweighted according to their distance from the central tendency of the sample. In each iteration, the quantity T_{bi}, representing the updated estimate of central tendency, is calculated, until the change in consecutive iterative values is negligible. The process is illustrated in detail in Appendix B.

As examples, random sets of 80 data points were selected from the calcium and ALT data described in Section 9.4. Each set was used to determine a 95% reference interval (for calcium in men, calcium in women, ALT in men, and ALT in women). As discussed above, the ALT data are skewed, but robust reference intervals may be derived in this case as well, even without transformation. (Outliers, if present, must still be removed, though, as discussed in Section 9.2.)

As reflected in Table 7, each of the resulting reference intervals using 80 observations is comparable to that obtained with the nonparametric method using all 120 observations. Although the robust method can be used with even smaller numbers of observations, the working group is hesitant to recommend that this be done, except in the most extreme instances. As noted earlier and as reflected in the next section, CIs are likely to be far too wide.

Table 7. Robust Reference Intervals Using Data Sets of $n = 80$

Method (Sample Size)	Calcium Women	Calcium Men	ALT Women	ALT Men
Robust $n = 80$	9.0-10.1	9.1-10.5	6-39	10-58
Nonparametric $n = 120$	8.9-10.2	9.2-10.3	6-46	10-55

9.5 Confidence Intervals for Reference Limits

The reference limits computed from a sample of selected subjects are estimates of the corresponding percentiles in the population of persons studied. Another sample of persons from the same population probably yields somewhat different reference limits. A useful way of recognizing and assessing the variability in sample estimates is by computing a CI for the population percentile being estimated, using information provided by the sample. In the present case, a CI is a range of values that includes the true percentile (eg, the 2.5th percentile of the population) with a specified probability, usually 90% or 95%. This probability is called the "confidence level" of the interval.

The concept of CIs rests on the presumption that a representative sample of observations (in the case of the subject of this document, reference individuals) has been drawn from some defined population. This implies that each member of the population is equally likely to be selected. This ideal is often not attained in practice. The most that can be expected is that the sample of reference individuals selected is, in fact, healthy persons from whom reference specimens are secured under the recommended preanalytical conditions. The reference individuals are at least randomly obtained from a described pool, eg, laboratory employees. Therefore, the basic assumptions for the validity of CIs are the observations are obtained independently of each other, and the reference sample is representative of the population even if not drawn strictly at random.

Confidence intervals are useful for two practical reasons. First, they remind the investigator of the variability of estimates and provide a quantitative measure of this variability. Second, CIs narrow as the size of the sampling increases. Therefore, an investigator may choose a larger sampling of reference individuals in order to obtain improved precision in the estimated reference interval.

9.5.1 Confidence Intervals for the Nonparametric Method

Nonparametric CIs are given by the observed values corresponding to certain rank numbers. Table 8[5,37] shows the rank number defining the 90% CI for the 2.5th percentile based on a given sample size. Note that the minimum sample size is 120; as noted earlier, although one can theoretically determine 95% reference intervals with as few as 39 samples, one needs a minimum of 120 samples to achieve 90% confidence limits on such intervals.

Table 8. Nonparametric Confidence Intervals of Reference Limits.[5] The table shows the rank numbers of the 0.90 confidence interval of the 0.025 fractile for samples with 119–1000 values. To obtain the corresponding rank numbers of the 0.975 fractile, subtract the rank numbers in the table from $N + 1$ where N is the sample size. (From: Solberg HE. Approved recommendations [1987] on the theory of reference values. Part 5. Statistical treatment of collected reference values. Determination of reference limits. *Journal of Clinical Chemistry and Clinical Biochemistry*. Vol. 25. Berlin, Germany: Walter de Gruyter GmbH & Co. KG; 1987, pp. 645-656. Table 1. Adapted with permission.)

Sample size	Rank numbers		Sample size	Rank numbers	
	Lower	Upper		Lower	Upper
119–132	1	7	566–574	8	22
133–160	1	8	575–598	9	22
161–187	1	9	599–624	9	23
188–189	2	9	625–631	10	23
190–218	2	10	632–665	10	24
219–248	2	11	666–674	10	25
249–249	2	12	675–698	11	25
250–279	3	12	699–724	11	26
280–307	3	13	725–732	12	26
308–309	4	13	733–765	12	27
310–340	4	14	766–773	12	28
341–363	4	15	774–799	13	28
364–372	5	15	800–822	13	29
373–403	5	16	823–833	14	29
404–417	5	17	834–867	14	30
418–435	6	17	868–871	14	31
436–468	6	18	872–901	15	31
469–470	6	19	902–919	15	32
471–500	7	19	920–935	16	32
501–522	7	20	936–967	16	33
523–533	8	20	968–970	17	33
534–565	8	21	971–1000	17	34

*ath lowest sample value = lower limit of 90% CI for 2.5th percentile in target population; bth lowest sample value = upper limit of 90% CI for 2.5th percentile in target population. To obtain ranks corresponding to a 90% CI for the 97.5th percentile, subtract the values given for a and b from $n + 1$.

For example, when the reference sample consists of 120 persons, the observations corresponding to rank numbers 1 and 7 define the 90% CI for the lower reference limits. To obtain the comparable rank numbers defining the 90% CI for the upper reference limit, these rank numbers are subtracted from 121 (in general, $n + 1$), giving 114 and 120. Thus, the smallest observation serves as the lower limit of the 90% CI for the lower reference limit, while the largest observation is the upper limit of the 90% CI for the upper reference limit.

Using the rank numbers in Table 8 and the data in Table 4 (calcium) and Table 5 (ALT), Table 9 presents the 90% CIs for the lower and upper reference limits for calcium and ALT.

Table 9. 90% Confidence Intervals for Lower and Upper 95% Reference Limits

Analyte	Lower Reference Limit	Upper Reference Limit
Calcium (mg/dL)*		
Women (n = 120)	8.8–9.1	10.1–10.3
Men (n = 120)	9.1–9.3	10.3–10.6
Combined (n = 240)	8.9–9.2	10.3–10.4
ALT (U/L)		
Women (n = 120)	5–8	36–65
Men (n = 120)	9–11	51–69
Combined (n = 240)	6–9	49–65

*mg/dL • 0.2495 = mmol/L

9.5.2 Confidence Intervals for the Reference Limits Obtained With the Robust Method

Confidence intervals for the reference limits obtained with the robust procedure cannot be calculated from a simple formula or by using statistical tables. Rather, they may be derived using a bootstrap resampling method. If the original dataset consists of *n* observations, then a resampling is achieved by sampling from these *n* observations with replacement, thus allowing some observations to appear multiple times while others may not appear at all. From this "pseudo" sample, the robust reference interval limits are derived as described earlier (see Section 9.4.2). This process is repeated a large number, B, of times. This yields B lower reference limits. The 5th and 95th percentiles of the B lower reference limit estimates constitute a 90% confidence interval for the lower reference limit. An analogous confidence interval is derived for the upper reference limit.[36,49]

Using this technique with the data from Tables 4 and 5, Table 10 indicates the 90% confidence limits for sample sizes 40 and 80 with the robust method for calcium and ALT in men. Note that, although the reference limits are comparable for $n = 40$ and $n = 80$, the confidence limits are much wider for $n = 40$ than for $n = 80$ (as they would be for any technique). For example, for $n = 40$, the confidence limits for the lower limit for calcium (9.0-10.1) are almost as wide as the reference interval itself (9.1-10.5). In contrast, for $n = 80$, the confidence limits span only 0.2 mg/dL (9.0-9.2), exactly the same as that for the nonparametric technique with $n = 120$. Similarly, for $n = 40$, the confidence limits for the upper limit for ALT are unacceptably high.

As noted earlier (see Section 9.1), many factors influence the width of the CIs. In this case, one could argue that, even for $n = 80$ with the robust method, the width of the CI for the upper limit of ALT is unacceptably large (51-64, a span of 13, 27% of the width of the reference interval, 10-58). It is, however, no worse than that determined with the nonparametric method with $n = 120$ (51-69, a span of 18, 40% of the width of the reference interval, 10-55). In this case, one would infer that more reference values are needed.

Table 10. Effect of Sample Size on Confidence Intervals

	Calcium in Males			ALT in Males		
Method (Sample Size)	**Reference Interval**	**Lower Limit CI**	**Upper Limit CI**	**Reference Interval**	**Lower Limit CI**	**Upper Limit CI**
Robust ($n = 40$)	9.1-10.5	9.0-10.1	10.3-10.7	9-60	7-12	45-71
Robust ($n = 80$)	9.1-10.5	9.0-9.2	10.4-10.6	10-58	8-11	51-64
Nonparametric ($n = 120$)	9.2-10.3	9.1-9.3	10.3-10.6	10-55	9-11	51-69

10 Transference

Because the determination of reliable reference intervals can be a major and costly task, it is very useful to be able to transfer a reference interval from one laboratory to another by some process less costly and more convenient. As more new tests and methods are introduced in more laboratories, it is unrealistic to expect each laboratory, large and small, to develop its own reference intervals. Consequently, clinical laboratories may rely more and more on other laboratories or diagnostic test manufacturers to generate and provide appropriate and adequate reference value data that can be transferred.

The transference of reference values requires that certain conditions be fulfilled in order to be acceptable. Assuming the original reference value study was done properly, the transference of the respective reference interval involves two distinct issues mentioned earlier (see Section 6.2):

(1) the comparability of the analytical system; and
(2) the comparability of the test subject population.

10.1 Transference: Comparability of the Analytical System

In some cases, laboratories have established, using the protocols described in this document, a reference interval for a specific analyte, with a specific methodology, for their subject population. If the laboratory decides to change methods, it may not be necessary to collect samples from reference individuals to establish a reference interval for the new method.

Rather, as part of its implementation of the new method, the laboratory has presumably undertaken a method comparison study between the two methods (see CLSI/NCCLS document EP09[8]). Data from such a study can be evaluated to see whether they can be used to update the existing reference interval or whether, in fact, a new study is needed.

Clearly, the major advantage of this strategy is it may obviate the need for the laboratory to obtain samples from reference individuals. One can use fresh samples from any patients (ie, not necessarily reference individuals) to investigate the relationship between the methods. In general, if the new method has similar imprecision and known interferences, uses the same or comparable standards or calibrators, and provides values that are acceptably comparable, then the reference interval can be transferred. Two examples are given below.

Example 1: Assays are completely comparable.

Consider the data from the current edition of CLSI/NCCLS document EP09.[8] As shown in Appendix B of that document, using the range of data selected (roughly 50 to 250) and linear regression, the best fit regression line is:

$$y = 1.004x - 0.628, r^2 = 0.990.$$

In this case, given the extremely large correlation coefficient, the small positive slope bias, the small negative intercept, and the range of values studied, the values from the two methods are comparable.

Assume the reference interval for the current method is 50 to 150. Based on the equation above, the new reference interval is still 50 to 150:

$$50 \rightarrow 50 \bullet 1.004 - 0.628 = 50.2 - 0.628 = 49.57, \text{ which rounds to } 50.$$
$$150 \rightarrow 150 \bullet 1.004 - 0.628 = 150.6 - 0.628 = 149.97, \text{ which rounds to } 150.$$

Example 2: Assay results are highly correlated, but one assay gives results that are proportionally biased higher or lower (eg, measuring enzyme activity at 37 °C instead of 30 °C).

Assume the new regression line is:

$$y = 1.57x - 0.832, r^2 = 0.990.$$

Even though the positive slope bias is now quite large, the correlation coefficient remains extremely high, and the intercept remains small. Thus, the reference range transforms to 78 to 235:

50 → 50 • 1.571 – 0.832 = 78.55 – 0.832 = 77.72, which rounds to 78.
150 → 150 • 1.571 – 0.832 = 235.65 – 0.832 = 234.82, which rounds to 235.

One should note some important caveats to using transference to determine a reference interval:

(1) One should follow the protocol in CLSI/NCCLS document EP09 very closely.[8] It is especially important to use an appropriate distribution of values. If there is not a sufficient range represented, the correlation may look worse than it is. On the other hand, if too large a range of samples is used (ie, many extreme low and extreme high values), one can overestimate the quality of the correlation.

(2) When using linear regression, it is critical to look at the magnitude of the intercept in comparison to the range of data and the reference interval. If the intercept is relatively large, compared to the reference interval, then method bias may negate a simple transfer. In this case, the working group recommends that the laboratory use samples from reference individuals to establish the reference interval.

(3) Linear regression may not always be the best, or most appropriate, tool to use for comparing the two sets of data. For example, sodium has a narrow range of values and these are discrete integers. In this case, simple difference of means may allow the reference interval to be set by correction of the mean bias between methods.

Even if one is able to use transference to calculate the new reference interval, the working group strongly encourages laboratories to validate the reference interval with a small sample (eg, n = 20) of reference individuals using the methods described in Section 11.2 below.

The working group does, however, also note that if the reference interval is validated on one subcategory (partition), there may not be a need to validate it on all the other subcategories (partitions).

10.2 Transference: Comparability of the Test Subject Population

If a clinical laboratory wishes to transfer a reference interval established by another laboratory or diagnostic test manufacturer for the same or acceptably comparable analytical system (as described in Section 9.1 above), the question of transference becomes one of comparability of the reference population. In addition, other preanalytical factors within the reference value study must also be comparable, such as preparation of the reference individuals and specimen collection and handling procedures. This type of transfer is increasingly common and probably accounts for most of the present reference interval assignments in clinical laboratories.

In this case, the working group again recommends validation by one of the three methods described in the next section.

11 Validation

Essentially, three approaches can be used to assess the acceptability of the transference of a reference interval:

(1) a subjective assessment;
(2) a statistical test on a relatively small number of reference individuals (eg, n = 20); and
(3) an evaluation of a larger number of reference individuals (but fewer than n = 120, the number needed to perform a standard reference interval study).

Each is described in detail below.

11.1 Validation: Subjective

The acceptability of the transfer may be rather subjectively assessed by a careful inspection of the pertinent factors of the original appropriate reference value study. To be able to do this, all of the reference population demographic variables and geographic locations must be adequately described and be available for review. Also, the preanalytical and the analytical procedural details, analytical performance, the complete set of reference values, and the method of estimating the reference interval must be stated. If, in the judgment of the laboratorian, these factors are consistent with the receiving laboratory's operation and test subject population, then the reference interval may be transferred without a requirement for any receiving laboratory validation studies, other than a documentation of these considerations.

11.2 Validation: Using Small Numbers of Reference Individuals

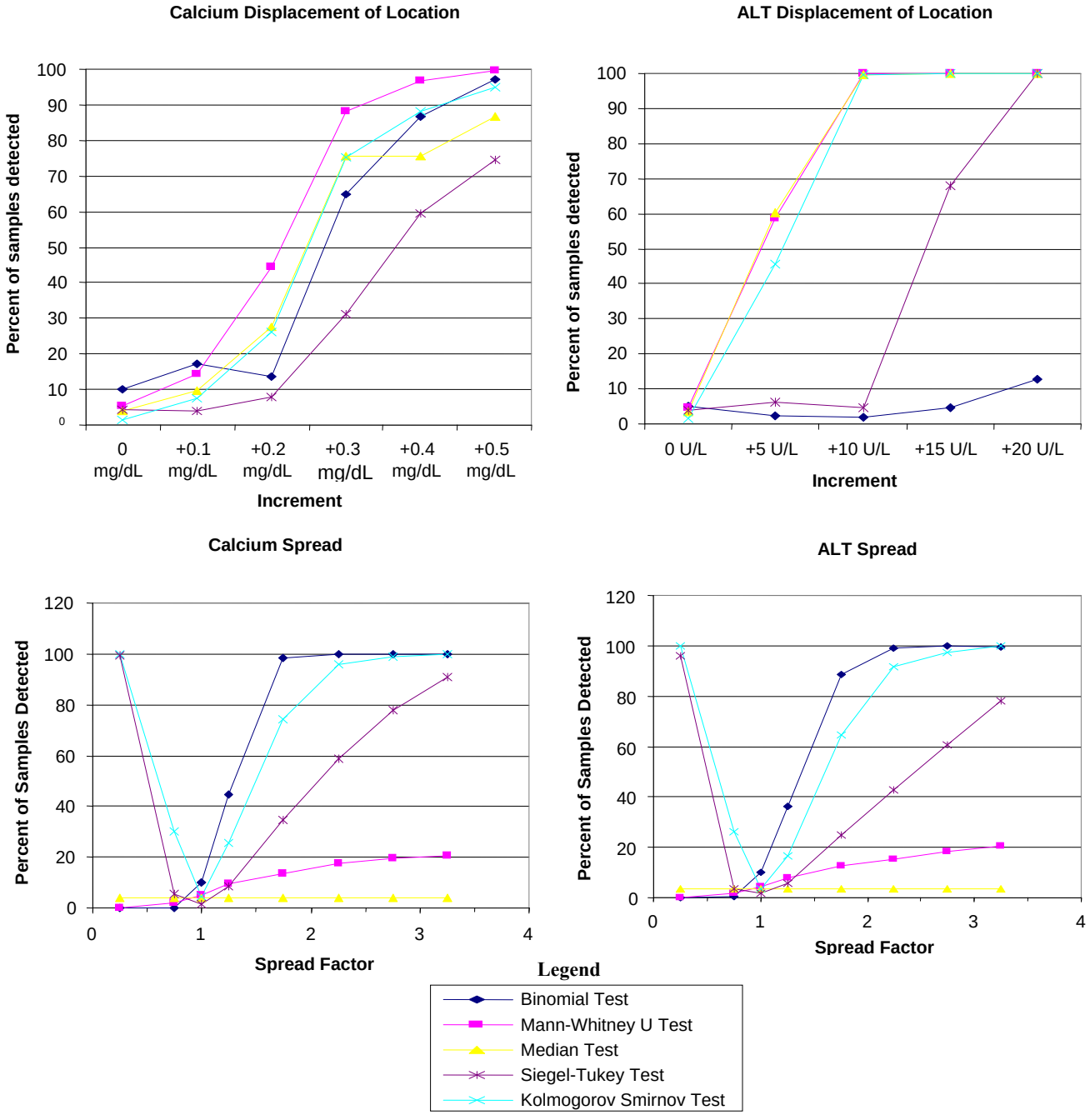
Alternatively, a user or receiving laboratory may wish to, or may be required to, validate the transference of a reference interval reported by a manufacturer or other donor laboratory. The acceptability of the transfer may be assessed by examining a smaller number of reference individuals (n = 20) from the receiving laboratory's own subject population and comparing these reference values to the larger, more adequate original study. Here again, however, the analytical and preanalytical factors of the original reference value study need to be consistent with the receiving laboratory's operation. Also, if there are substantial differences in the geographic locations or demographic variables of the two populations that are known to cause differences in the reference values, there is little point in trying to transfer the reference interval.

For the transference validation study, the reference individuals are selected and the reference values are obtained in accordance with the previously discussed guidelines. These 20 persons should reasonably represent the receiving laboratory's healthy population and satisfy the exclusion and partition criteria appropriately. After testing these 20 specimens according to the appropriate specifications, the test results should be examined to make sure they represent a statistically homogeneous group of results, ie, that none of the results appears to be an outlier. To test for possible outliers, the Reed/Dixon or Tukey methods cited earlier should be applied. Genuine outliers identified by these techniques should be eliminated and new patient specimens obtained to replace them, so 20 test results with no outliers are finally secured.

The manufacturer's or donor laboratory's reported 95% reference limits may be considered valid for application in the receiving laboratory if no more than two of the 20 tested subjects' values (or 10% of the test results) fall outside those original reported limits. If three or four test results fall outside these limits, another 20 reference specimens similar to the first 20 should be obtained, again free of outliers. If no more than two of these new results fall outside the manufacturer's or donor laboratory's reported reference limits, the latter may be considered acceptable for use in the receiving laboratory. However, if three or more again fall outside the limits (or if five or more in the original set fall outside the limits), the user should reexamine the analytical procedures used, consider possible differences in the biological characteristics of the two populations sampled, and consider whether the receiving laboratory should develop its own reference interval according to the full-scale study guidelines.

This approach, calling for the receiving laboratory to test 20 selected subjects using the comparable or same method of analysis, and accepting the manufacturer's or donor laboratory's limits if two or fewer test results fall outside those limits, is statistically sound, as may be proven by recourse to tables of the binomial distribution. The probability that more than two test results will fall outside those limits, when, in fact, 95% of the user's population falls within those limits, is only 7.5%. When one considers the rule in its entirety (ie, collecting samples from an additional 20 reference individuals when three or four values in the original set were outside the proposed limits), the probability of false rejection drops to just under 1%.

In those cases where the full dataset from tested reference individuals is available from the assay manufacturer or donor laboratory, other statistical tests, potentially more powerful than the binomial test just described, can be undertaken. The working group performed computer simulations on the calcium and ALT data by introducing changes of various magnitudes in location (mean) and in scale (spread). Figure 4 shows the relative power of several statistical procedures for detecting differences in distribution between the receiving laboratory's reference individuals and the original laboratory's proposed reference interval. See Appendix A for details on the simulation.



NOTE: In the figures above, 1000 random samples of 20 values each were selected for each analyte at each x-axis value. For displacement of location, the increment listed on the x-axis was added to each value. For spread, each value's difference from the median was multiplied by the spread factor on the x-axis. Then, the various statistical tests were applied to see how frequently the resulting change was detected.

Figure 4. Performance Characteristics of Validation Tests

The Mann-Whitney U (M-W U) test is very powerful in detecting changes in location. At virtually every increment tested, it detected the changes at the highest rate. For example, it detected a calcium increment of 0.2 mg/dL almost 50% of the time (vs 15% for the binomial test), and it detected an ALT increment of 5 U/L roughly 60% of the time (vs 0% for the binomial test).

The Siegel-Tukey (S-T) test for differences in scale is a very powerful method for detecting increases and decreases in scale. For example, it detected a 25% increase in scale roughly 25% of the time for both calcium and ALT (vs roughly 40% for the binomial test), and it detected a 25% decrease in scale roughly 25% of the time for both calcium and ALT (vs 0% for the binomial test).

Overall, one can infer that the binomial test performs reasonably well in detecting changes in location of symmetrically distributed variables such as calcium, but poorly in detecting such changes in markedly skewed distributions such as ALT. Similarly, the binomial test performs reasonably well in detecting increases in scale, but it does not detect decreases in spread at all. Thus, when using the binomial test to validate a reference interval, one should be particularly sensitive to the possibility that the proposed reference interval may be too wide for the target population, especially when none of the 20 sample values falls outside the limits. This situation could arise, for instance, if the target population is more homogeneous than the original population (eg, men between the ages of 20 and 30 vs both sexes between the ages of 20 and 70), or if the method itself is more precise than the original method. As was noted at the beginning of this section, it is critical that comparability of methods and of populations be established prior to undertaking a validation study.

With respect to the other statistical tests, the M-W U test is best at detecting changes in location, the S-T test is best at detecting changes in spread, and the Kolmogorov Smirnov (K-S) test is probably the best single test overall to detect changes in location and spread (both increase and decrease).

11.3 Validation: Using Larger Numbers of Reference Individuals

Laboratories may elect to undertake a more extensive reference interval transference study for analytes whose reference intervals are critically important for local clinical interpretation of the assay. In such cases, the acceptability of the transfer may be assessed and validated by examining a larger population of reference individuals (eg, $n = 60$) from the receiving laboratory's own subject population and comparing these reference values to the larger, more adequate original study. Here again, however, the analytical and preanalytical factors of the original reference value study need to be consistent with the receiving laboratory's operation. Larger studies of this sort have more statistical power for discovering differences between the original reference value study and the receiving laboratory subject population. If substantial differences exist in the geographic locations or demographics of the two populations that are known to cause differences in the reference values, there is little point in trying to transfer the reference interval.

For such studies, the reference individuals are selected and the reference values are obtained in accordance with the previously discussed guidelines in Sections 6 and 7. After appropriate examination of the data and the exclusion of any outliers, the smaller sample of reference values is compared with the larger original set of reference values from the donor laboratory.

The two sets of reference values may be treated in the same manner as described for determining whether subclasses exist in a reference population (see Section 9.3). If this evaluation does not demonstrate a large, significant difference (a subclass difference) between the donor reference values and the receiving laboratory's briefer set of reference values, the donor reference interval may be transferred. However, if the difference is significant according to the partitioning protocol, further comparison or a full-scale reference value study should be undertaken.

Again, the availability of robust statistical techniques provides another alternative. As suggested in Section 9.5.2, with a sample of 60 reference individuals, robust techniques may allow reference intervals

with reasonably narrow confidence limits to be established. However, each laboratory needs to decide, on an analyte-by-analyte basis, whether the confidence limits on such reference intervals are sufficiently narrow to meet clinical interpretational needs.

12 Presentation of Reference Values

12.1 Introduction

This section addresses the presentation of patient values related to reference values. The comments in this section are divided into two groups. Section 12.2 addresses the presentation of reference values by laboratories and end users. Section 12.3 covers the same subject as it applies to the manufacturers of quantitative clinical laboratory diagnostics tests.

12.2 Laboratory Presentation

Every quantitative clinical result should be accompanied by an appropriately presented reference interval. The reference intervals applied should reflect the subclass partitions that are determined to be significant for that laboratory's particular reference population. Lengthy reports that include the results of many tests should include some way of highlighting those results not within the reference interval (for example, printing "high" or "low" adjacent to a result). The term "reference interval" should be used; the terms "normal," "usual," or "expected" should be avoided. Figure 5 shows an example of an acceptable report.

The use of forms with preprinted reference intervals requires that reference intervals for all appropriate subclasses be included and, as a result, may prove to be confusing. A better approach is for the computer or instrument to print the reference interval appropriate for the particular patient. In most cases, the subclass reference intervals are determined by the age and sex of the patient. Any report that uses subclass reference intervals should have the patient's partitioning factors included in the header or the demographics portion of the report.

The origin of the reference intervals used should be documented. If they are obtained internally, all the documentation (number and demographics of the reference individuals, the assessment of health criteria used, the exclusion and partitioning criteria used with the reference sample, size of the subclasses, analytical and preanalytical details) should be maintained and be made available on request. (If the reference intervals are verified by some other method, the supporting data should likewise be maintained and be made available on request.)

Whenever changes are made in reference intervals, a separate communication regarding the changes should be sent to all users of the laboratory, as well as indicated on the report.

12.2.1 Medical Decision Limits

As described in this document, reference intervals are defined by statistical methods and are descriptive of a specific population. In contrast, decision limits are defined by consensus and distinguish among different populations. When decision limits determined by national or worldwide consensus exist, these limits, rather than reference intervals, should be reported.

As examples, consider high-density lipoprotein (HDL) cholesterol and NTproBNP/BNP. In the case of HDL cholesterol, decision limits can be used to categorize people as having increased risk (<40 mg/dL) or decreased risk (>60 mg/dL) for coronary artery disease based on data from large population studies.[50] In the case of BNP/NTproBNP, decision limits, based on clinical sensitivity and clinical specificity, can be used to determine the likelihood of a patient having congestive heart failure.[51,52]

To avoid confusion, the working group encourages laboratories to report either decision limits or reference intervals but not both, with a clear indication of which has been used (eg, total cholesterol and HDL cholesterol in the sample report in Figure 5). When several decision limits are reported for different clinical situations (eg, for low-density lipoprotein [LDL] cholesterol), a brief summary of the recommendations should be added to the report.

The Hospital Laboratory
1440 Main Street
Anywhere, State 12345

Patient's Name	Smith, James	**Date/Time Collected**	07/19/2007 5:35 am
Medical Record #	XXXX		
Age	57	**Sample ID**	XXXXX
Sex	Male	**Sample Comment**	Fasting
		Collected By	JLF
Ordering Physician	XXXXX		

Test Name	**Results**		**Units**	**Reference Interval**
Sodium	140		mmol/L	133-145
Potassium	4.5		mmol/L	3.3-5.1
Chloride	105		mmol/L	96-108
Bicarbonate	28		mmol/L	22-32
ALT	43		U/L	10-55
AST	32		U/L	8-45
Total Cholesterol	5.85*	HIGH	mmol/L	<5.17†
HDL Cholesterol	0.98*	LOW	mmol/L	>1.04†
Calcium	2.68	HIGH	mmol/L	2.28-2.58
Phosphorus	1.04		mmol/L	0.87-1.45

* mmol/L● (38.67) = mg/dL
† National Cholesterol Education Program (NCEP) recommendations

Figure 5. Sample Laboratory Report

12.3 Manufacturer Presentation

Manufacturers of quantitative diagnostic tests should provide detailed reference interval information in their product labeling. For tests that are well studied and have widely recognized factors that partition reference samples into subclasses, manufacturers should provide reference intervals for such subclasses. They should indicate whether the most common partitioning factors are examined for subclass differences (eg, sex, age, fasting/nonfasting, time of day, pregnancy, and posture). It is important to recognize that there may be subclass differences in reference individuals from region to region that reflect not only geographical differences, but also other variables such as environment, diet, and ethnic background.

In all cases, manufacturers should make use of guidelines in this document, including information regarding control of preanalytical and analytical variables, as well as enumerating, at a minimum, the following data about the sample population used:

- number;
- sex;
- age (as appropriate);
- ethnicity and/or race (if relevant); and
- percentiles used for cutoffs (eg, central 95%, 99th percentile).

In all cases (as mentioned in Section 8.3), the working group encourages manufacturers to use methods able to provide results traceable per ISO Standard 17511[9] to internationally recognized standards, so as to facilitate comparisons of patient values between methods.

When reference intervals are established, it is important that the sponsors describe them in sufficient detail. They should define the analytical methods used, the reference individuals studied, and the technique by which the reference limits are chosen. For example:

> *Using reagents X, standards Y, and instrumentation Z, we measured Analyte A on 120 apparently healthy men, aged 20 to 30. No outliers were detected by the Tukey method (1977). Using the nonparametric technique, the two highest and the two lowest values were then eliminated. The resulting reference interval includes the remaining 95% of samples.*

Manufacturers of laboratory equipment, especially of data management systems, should provide the capability of printing the reference intervals for subclasses, as well as the associated patient demographics as described in Section 12.2.

When manufacturers make substantial modifications to the assay system (reagents, calibrators, and/or instrumentation), they should provide explicit documentation regarding effects on the reference interval.

13 Other Issues

13.1 Qualitative Analysis

The evaluation of reference data generated from qualitative analyses is not within the scope of this document.

13.2 Therapeutic Drug Levels

This guideline does not address the determination of therapeutic drug levels. This requires a different study. The population required for these studies necessarily has to be under the influence of the pharmacologic agent and at a clinically effective level. This problem is complex and involves a number of additional issues such as dosage, dosing, time of specimen procurement in relation to time of administration of the drug, the route of drug administration, clinical effectiveness, toxicity, and other issues.[53]

13.3 Time-Dependent/Challenge Tests

It is beyond the scope of this document to provide the user with all the necessary details to set up protocols for time-dependent and challenge tests or studies that require serial measurements. Clearly, there are many other factors to consider in addition to all those that are of "routine" concern.

13.4 Individual Variation

This document deals with population-based reference intervals only and does not address the issue of "individual" reference intervals where the individual subject is the referent. This involves a separate study of the biological component for the total variance of observed values in each subject under given experimental conditions.[54]

13.5 "Critical Values"

This guideline is not intended to address the issue of "critical values," those values that require immediate notification of the ordering physician. These values should be determined by each laboratory in consultation with its respective physicians, as practices may vary widely among institutions.

14 Summary

In this document, the working group strongly endorses the approach and systematic process for determining reference intervals described in the previous version. Certain aspects of that process have received new emphasis, and a few new concepts have been added. The working group believes the process remains reasonable while providing a reliable foundation for the production of reliable reference intervals. The basic principles that follow are uniformly important and must underlie any reference value study:

(1) The selection of reference individuals must be thoughtful, with advance consideration given to exclusion and partitioning criteria. The reference population must be appropriate and useful to the process of determining disease or abnormalities in the patient population. The evaluation of the health status of the reference individuals must be documented and described as part of the reference value study or reference intervals defined. The better the reference individuals are defined and described, the greater the value of the reference interval studies.

a) The working group again rejected the concept of a "gold standard" reference population of absolutely healthy young adults as a prerequisite for the determination of a health-associated reference interval.

b) As a general rule, the use of hospital or clinic patients as a source for reference individuals was also rejected. Patient data should only be used for deriving a reference interval when "nonpatient" reference individuals cannot be obtained, and only with careful selection and attention to exclusion and partitioning criteria.

(2) All of the preanalytical and analytical processes related to the measurement of reference values must be thoughtfully considered and controlled where appropriate. It is essential that these factors be treated in the same manner for the reference individuals as for the patient population tested.

(3) Once the data are collected, a frequency histogram should be prepared and examined visually in order to facilitate analysis. A process for detecting and discarding outlier values is recommended. In addition to the Dixon-Reed rule recommended in the previous edition of C28, an alternative rule, based on Tukey,[41] was added to this document.

(4) The nonparametric method of estimation of the reference interval is again strongly recommended as the preferred method for analysis because of its simplicity and reliability. More importantly, this method requires no specific assumption about the mathematical form of the probability distribution of reference values.

For the nonparametric method, a minimum sample of 120 reference values is recommended for each reference population or subclass. This is the smallest number of samples that allows the determination of a 90% CI around the reference limits (eg, the 2.5th and 97.5th percentiles). Greater confidence or improved precision in an estimated 95% reference interval can be accomplished using a larger sample of reference individuals.

(5) Recognizing the difficulty for individual laboratories to obtain sufficient numbers of reference individuals, the working group has introduced the concepts of multicenter trials and robust statistical methods.

a) Well-organized multicenter trials should allow for pooling of data from multiple sites. By ensuring comparability in analytical methods as well as adhering to strict selection criteria, the only remaining reasons precluding pooling of data might relate to population differences such as race and region.

b) Even in the absence of multicenter trials, individual laboratories may be able to establish reference intervals with smaller numbers of reference individuals by employing modern statistical methods. Examples of one of those techniques are provided.

(6) The working group has placed new emphasis on the concept of confidence limits of reference intervals. With too few points, confidence limits can be so wide as to make the reference intervals virtually meaningless. As noted, even though one can theoretically establish 95% reference intervals with the nonparametric method using just 39 values, one actually needs 120 values to obtain 90% confidence limits for such intervals. For any method of data analysis, the use of more points translates into tighter, and more useful, confidence limits.

(7) A rigorous and systematic approach is recommended for determining when separate reference intervals for subclasses are necessary.

(8) In those cases where a laboratory implementing a new analytical method wants to adapt a reference interval determined previously on its own patient population, the process of transference can be used. Several specific caveats are described. In addition, the working group strongly encourages laboratories to verify the new reference interval with a small group of reference individuals.

(9) The working group recognizes that **establishing** reference intervals is beyond the capability of most individual laboratories. However, the working group believes that **verifying** reference intervals established elsewhere (eg, manufacturers' product inserts) is feasible for most individual laboratories.

One can, with as few as 20 samples from reference individuals, use a relatively simple test to verify the applicability of a reference interval to one's own population. The performance characteristics of this test and several other tests are described.

(10) In increasing numbers of cases (eg, cholesterol, glycated hemoglobin), establishing and verifying traditional reference intervals as described in this document is not appropriate. For such analytes, where national (or international) consensus on decision limits exists, it is critical that manufacturers and laboratories ensure their methods provide accurate results on patient samples.

(11) Recommendations regarding presentation of reference intervals are made for both manufacturers and laboratories.

References

1 Solberg HE. Approved recommendation (1986) on the theory of reference values. Part 1. The concept of reference values. *Clin Chim Acta.* 1987;167:111-118.

2 PetitClerc C, Solberg HE. Approved recommendation (1987) on the theory of reference values. Part 2. Selection of individuals for the production of reference values. *J Clin Chem Clin Biochem.* 1987;25:639-644.

3 Solberg HE, PetitClerc C. Approved recommendation (1988) on the theory of reference values. Part 3. Preparation of individuals and collection of specimens for the production of reference values. *Clin Chim Acta.* 1988;177:S1-S12.

4 Solberg HE, Stamm D. Approved recommendation on the theory of reference values. Part 4. Control of analytical variation in the production, transfer, and application of reference values. *Eur J Clin Chem Clin Biochem.* 1991;29:531-535.

5 Solberg HE. Approved recommendations (1987) on the theory of reference values. Part 5. Statistical treatment of collected reference values. Determination of reference limits. *J Clin Chem Clin Biochem.* 1987;25:645-656.

6 Dybkaer R, Solberg HE. Approved recommendations (1987) on the theory of reference values. Part 6. Presentation of observed values related to reference values. *J Clin Chem Clin Biochem.* 1987;25:657-662.

7 CLSI. *Metrological Traceability and Its Implementation; A Report*. CLSI document X05-R. Wayne, PA: Clinical and Laboratory Standards Institute; 2006.

8 CLSI/NCCLS. *Method Comparison and Bias Estimation Using Patient Samples; Approved Guideline—Second Edition.* CLSI/NCCLS document EP09-A2. Wayne, PA: NCCLS; 2002.

9 ISO. *In vitro* diagnostic medical devices — Measurement of quantities in biological samples — Metrological traceability of values assigned to calibrators and control materials. ISO 17511. Geneva: International Organization for Standardization; 2003.

10 Guideline for Isolation Precautions: Preventing Transmission of Infectious Agents in Healthcare Settings 2007. Available at: http://www.cdc.gov/ncidod/dhqp/gl_isolation.html. Accessed 20 September 2008.

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

12 ISO. *Statistics – Vocabulary and symbols – Part 1: General statistical terms and terms used in probability*. ISO 3534-1. Geneva: International Organization for Standardization; 2006.

13 Levey AS, Bosch JP, Breyer Lewis J, Greene T, Rogers N, Roth D, for the Modification of Diet in Renal Disease Study Group. A more accurate method to estimate glomerular filtration rate from serum creatinine: A new prediction equation. *Ann Intern Med*. 1999;130:461-470.

14 Myron Johnson A, Hyltoft Petersen P, Whicher JT, Carlström A, MacLennan S, on behalf of the International Federation of Clinical Chemistry and Laboratory Medicine Committee on Plasma Proteins. Reference intervals for serum proteins: similarities and differences between adult Caucasian and Asian Indian males in Yorkshire, UK. *Clin Chem Lab Med*. 2004;42:792-799.

15 Baadenhuijsen H, Smit JC. Indirect estimation of clinical chemical reference intervals from total hospital patient data: application of a modified Bhattacharya procedure. *J Clin Chem Clin Biochem*. 1985;23:829-839.

16 Bock BJ, Dolan CT, Miller GC, et al. The data warehouse as a foundation for population-based reference intervals. *Am J Clin Pathol*. 2003;120:662-670.

17 Ferre-Masferrer M, Fuentes-Arderiu X, Puchal-Ane R. Indirect reference limits estimated from patients' results by three mathematical procedures. *Clin Chim Acta*. 1999;279:97-105.

18 Grossi E, Colombo R, Cavuto S, Franzini C. The REALAB project: a new method for the formulation of reference intervals based on current data. *Clin Chem*. 2005;51:1232-1240.

19 Kouri T, Kairisto V, Virtanen A, et al. Reference intervals developed from data for hospitalized patients: computerized method based on combination of laboratory and diagnostic data. *Clin Chem*. 1994;40:2209-2215.

20 Hoffman RG. Statistics in the practice of medicine. *JAMA.* 1963;185:864-873.

21 Statland BE, Winkel P. Selected pre-analytical sources of variation in reference values. In: Grasbeck R, Alstrom T, eds. *Reference Values in Laboratory Medicine.* Chichester, England: John Wiley and Sons Ltd.; 1981:127-137.

22 Solberg HE. Establishment and use of reference values. In: Tietz NW, ed. *Textbook of Clinical Chemistry*. Philadelphia, PA: WB Saunders; 1986:356-386.

23 CLSI. *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Sixth Edition*. CLSI document H03-A6. Wayne, PA: Clinical and Laboratory Standards Institute; 2007.

24 CLSI. *Procedures and Devices for the Collection of Diagnostic Capillary Blood Specimens; Approved Standard—Sixth Edition*. CLSI document H04-A6. Wayne, PA: Clinical and Laboratory Standards Institute; 2008.

[25] CLSI/NCCLS. *Procedures for the Collection of Arterial Blood Specimens; Approved Standard—Fourth Edition*. CLSI/NCCLS document H11-A4. Wayne, PA: NCCLS; 2004.

[26] CLSI. *Collection, Transport, and Processing of Blood Specimens for Testing Plasma-Based Coagulation Assays and Molecular Hemostasis Assays; Approved Guideline—Fifth Edition*. CLSI document H21-A5. Wayne, PA: Clinical and Laboratory Standards Institute; 2008.

[27] CLSI/NCCLS. *Urinalysis and Collection, Transportation, and Preservation of Urine Specimens; Approved Guideline—Second Edition*. CLSI/NCCLS document GP16-A2. Wayne, PA: NCCLS; 2001.

[28] CLSI. *Analysis of Body Fluids in Clinical Chemistry; Approved Guideline*. CLSI document C49-A. Wayne, PA: Clinical and Laboratory Standards Institute; 2007.

[29] CLSI/NCCLS. *Procedures for the Handling and Processing of Blood Specimens; Approved Guideline—Third Edition*. CLSI/NCCLS document H18-A3. Wayne, PA: NCCLS; 2004.

[30] CLSI. *Statistical Quality Control for Quantitative Measurement Procedures: Principles and Definitions; Approved Guideline—Third Edition*. CLSI document C24-A3. Wayne, PA: Clinical and Laboratory Standards Institute; 2006.

[31] Stamm D. Control of analytical variation in the production of reference values. In: Grasbeck R, Alstrom T, eds. *Reference Values in Laboratory Medicine.* Chichester, UK: John Wiley and Sons Ltd.; 1981:109-126.

[32] Burtis CA, Ashwood ER, Bruns DE, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics.* 4th ed. St. Louis, MO: Elsevier Saunders; 2006:425-448.

[33] Harris EK, Boyd JC. *Statistical Basis of Reference Values in Laboratory Medicine.* New York: Marcel Dekker; 1995.

[34] Linnet K. Nonparametric estimation of reference intervals by simple and bootstrap-based procedures. *Clin Chem.* 2000;31:867-869.

[35] Shultz EK, Willard KE, Rich SS, Connelly DP, Critchfield GC. Improved reference-interval estimation. *Clin Chem.* 1985;31:1974-1978.

[36] Horn PS, Pesce AJ. *Reference Intervals. A User's Guide*. Washington, DC: AACC Press; 2005.

[37] Reed AH, Henry RJ, Mason WB. Influence of statistical method used on the resulting estimate of normal range. *Clin Chem.* 1971;17:275-284.

[38] Linnet K. Two-stage transformation systems for normalization of reference distributions evaluated. *Clin Chem.* 1987;33:381-386.

[39] Barnett V, Lewis T. *Outliers in Statistical Data.* Chichester, England: John Wiley and Sons, Ltd.; 1978:68-73.

[40] Dixon WJ. Processing data for outliers. *Biometrics.* 1953;9:74-89.

[41] Tukey JW. *Exploratory Data Analysis*. Reading, MA: Addison-Wesley; 1977:688.

[42] Box G, Cox D. An analysis of transformations. *J Royal Statistical Society.* 1964;B26:211-252.

[43] Sinton TJ, Crowley D, Bryant SJ. Reference values for calcium, phosphate, and alkaline phosphatase as derived on the basis of multichannel-analyzer profiles. *Clin Chem.* 1986;32:76-79.

[44] Harris EK, Boyd JC. On dividing reference data into subgroups to produce separate reference ranges. *Clin Chem.* 1990;36:265-270.

[45] Lahti A. Partitioning biochemical reference data into subgroups: comparison of existing methods. *Clin Chem Lab Med.* 2004;42:725-733.

[46] Lahti A, Petersen PH, Boyd JC, Rustad P, Laake P, Solberg HE. Partitioning of nongaussian-distributed biochemical reference data into subgroups. *Clin Chem.* 2004;50:891-900.

[47] Lahti A. Are the common reference intervals truly common? Case studies on stratifying biochemical reference data by countries using two partitioning methods. *Scand J Clin Lab Invest*. 2004;64:407-430.

[48] Harris EK, Wong ET, Shaw ST. Statistical criteria for separate reference intervals: race and gender groups in creatine kinase. *Clin Chem.* 1991;37:1580-1582.

[49] Efron B. *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia, PA: Society for Industrial and Applied Mathematics; 1982.

[50] Executive summary of the third report of the National Cholesterol Education Program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol adults (Adult Treatment Panel III). *JAMA.* 2001;285:2486-2497.

[51] Januzzi JL, Camargo CA, Anwaruddin S, et al. The N-terminal Pro-BNP investigation of dyspnea in the emergency department (PRIDE) study. *Am J Cardiol.* 2005;95:948-954.

[52] Maisel AS, Krishnaswamy P, Nowak R, et al. Rapid measurement of B-type natriuretic peptide in the emergency diagnosis of heart failure. *N Engl J Med.* 2002;347:161-167.

[53] Bowers LD. Analytical goals in therapeutic drug monitoring. *Clin Chem.* 1998;44:375-380.

[54] Winkel P. The use of the subjects as his own referent. In: Grasbeck R, Alstrom T, eds. *Reference Values in Laboratory Medicine.* Chichester, England: John Wiley and Sons Ltd.; 1981:65-78.

Appendix A. Effectiveness of Several Statistical Tests in Validating Transference of Reference Intervals

When using 20 reference samples from one's own laboratory to evaluate the validity of transferring a reference interval from a manufacturer or donor laboratory, statistical tests other than the binomial test (described in Section 10.2) can be carried out. These tests, however, require that the full dataset from tested reference individuals be made available from the assay manufacturer or donor laboratory. Such tests include the Mann-Whitney U (M-W U) test (Wilcoxon Rank Sum test), the Median test, the Siegel-Tukey (S-T) test of scale (distributional spread), and the Kolmogorov Smirnov (K-S) test of overall fit to cumulative distribution. These statistical tests are available in standard statistical packages.

A1. Computer Simulation Studies

The working group undertook computer simulation studies to compare the powers of each of these methods to detect changes in location and scale. To perform these simulations, 597 results for each test from the men in five consecutive medical school classes constituted the reference population. In an experiment using random sampling with replacement, 1000 samples of 20 results each were selected from the 597 results for each test.

To simulate changes in location (displacement) of the results, each random sample was modified by adding a set increment to the result. Then, each of the above statistical tests was performed on the 20-result sample to derive the rates shown in the graphs in Figure 4.

To simulate changes in scale (spread) of the results, each random sample was modified according to the following transformations to expand (values of spread factor >1) or contract (values of spread factor <1) the test data scale centered on the median value of the distribution. Spread factors between 0.25 and 3.25 were utilized in these simulations. The medians of the 597 calcium and 597 ALT results in the reference population were 9.6 mg/dL (2.39 mmol/L) and 23 U/L, respectively.

$$\text{Calcium}_{\text{modified}} = [\text{Spread Factor} \bullet (\text{Calcium}_{\text{original}} - 9.6)] + 9.6$$

$$\text{ALT}_{\text{modified}} = [\text{Spread Factor} \bullet (\text{ALT}_{\text{original}} - 23)] + 23$$

Again, each of the statistical tests was performed on each sample to derive the rates shown in the graphs in Figure 4.

Appendix B. Robust Calculation

In order to illustrate the performance of the robust method, the data set chosen consists of 20 random observations from the female calcium data set in Table 4. As noted in Sections 9.4.2 and 9.5.2, the working group does not endorse such a small sample set for determining reference intervals, but its use in this example allows more efficient use of space while demonstrating a sufficient level of detail.

The data set is listed in the first column of the table below, sorted by increasing value to clarify the concepts underlying the calculation. As the initial estimate of location (center), T_{bi}, the median (9.6) is used; and for the initial estimate of scale (spread), the median absolute deviation about the median (MAD), in this case 0.10, is used, divided by 0.6745 (this factor is used to "standardize" the MAD so it is consistent with the standard deviation from a Gaussian distribution). The formula for weighting is, for each observation x_i:

$w_i = (1-u_i^2)^2$ *so long as* $-1 < u_i < 1$, *where* $u_i = (x_i - T_{bi})/(c \bullet MAD/0.6745)$

(Note that if $u_i < -1$ *or* $u_i > 1$, *then the weight is set equal to zero.)*

The variable *c* represents a tuning constant, which is set to 3.7 here (and 205.6 in a later calculation). At relatively low values like 3.7, the tuning constant makes the estimate of T_{bi} more resistant to outliers; at higher values like 205.6, the tuning constant helps to capture the variability in the population. These tuning constants were chosen to accommodate several distributions, including the Gaussian, as well as heavier-tailed distributions.[1]

T_{bi}, the estimate of location (center), is then calculated according to the formula:

$$T_{bi} = \frac{\sum w_i \bullet x_i}{\sum w_i}$$

The calculation of T_{bi} is repeated iteratively, using the updated values of T_{bi}, until the change in consecutive iterative values is negligible (< 0.001%).

Appendix B. (Continued)

For this example, the weights and T_{bi} over six iterations are as follows:

x_i		Weights iteration 1	Weights iteration 2	Weights iteration 3	Weights iteration 4	Weights iteration 5	Weights iteration 6
8.9		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9.2		0.2193	0.1798	0.1671	0.1630	0.1616	0.1612
9.4		0.7518	0.7132	0.6998	0.6953	0.6938	0.6933
9.4		0.7518	0.7132	0.6998	0.6953	0.6938	0.6933
9.5		0.9346	0.9121	0.9039	0.9010	0.9001	0.8998
9.5		0.9346	0.9121	0.9039	0.9010	0.9001	0.8998
9.5		0.9346	0.9121	0.9039	0.9010	0.9001	0.8998
9.6		1.0000	0.9982	0.9969	0.9963	0.9961	0.9960
9.6		1.0000	0.9982	0.9969	0.9963	0.9961	0.9960
9.6		1.0000	0.9982	0.9969	0.9963	0.9961	0.9960
9.6		1.0000	0.9982	0.9969	0.9963	0.9961	0.9960
9.7		0.9346	0.9540	0.9597	0.9616	0.9622	0.9624
9.7		0.9346	0.9540	0.9597	0.9616	0.9622	0.9624
9.7		0.9346	0.9540	0.9597	0.9616	0.9622	0.9624
9.7		0.9346	0.9540	0.9597	0.9616	0.9622	0.9624
9.7		0.9346	0.9540	0.9597	0.9616	0.9622	0.9624
9.8		0.7518	0.7883	0.8000	0.8039	0.8051	0.8056
9.9		0.4913	0.5366	0.5517	0.5567	0.5583	0.5589
9.9		0.4913	0.5366	0.5517	0.5567	0.5583	0.5589
10.2		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
T_{bi} =	9.6	9.6163	9.6217	9.6236	9.6242	9.6244	9.6244

As shown in the table, actual observations are down weighted according to their distance from the central tendency of the sample, with any observations more than 3.7 times the MAD/0.6745 getting zero weight. In this example, the values 8.9 and 10.2 receive weighting factors of zero, whereas the value of 9.6, the median, starts with a weighting factor of 1.0 and ends with a weighting factor of 0.9660.

With T_{bi} calculated, the 95% reference interval can be determined as:

$$T_{bi} \pm t_{n-1}^{0.025} \sqrt{s_{bi}^2(205.6) + S_T^2(3.7)}$$

where:

$t_{n-1}^{0.025}$ is the upper 2.5th percentile of a Student's t with (n-1) degrees of freedom and n is the sample size.

s_{bi} [205.6] is the biweight estimator of spread with tuning constant 205.6.

S_T [3.7] is the biweight estimator of the variability of $T_{bi.}$

The formulas for s_{bi} and S_T look formidable, but the values can be calculated in a relatively straightforward way using a standard spreadsheet program like Microsoft Excel. Note that in the formulas that follow, the construction max (x,y) means that one should use the maximum of the values of x and y.

Appendix B. (Continued)

$$s_{bi}[205.6] = 205.6 \times s \times \sqrt{\frac{n \sum_{-1<u_i<1} u_i^2 (1-u_i^2)^4}{\left(\sum_{-1<u_i<1} (1-u_i^2)(1-5u_i^2)\right) \times \max\left(1, -1 + \sum_{-1<u_i<1} (1-u_i^2)(1-5u_i^2)\right)}}$$

$$\text{where } u_i = \frac{x_i - median}{205.6 \times s};\ s = MAD/0.6745$$

$$s_{bi}[3.7] = 3.7 \times s \times \sqrt{\frac{n \sum_{-1<u_i<1} u_i^2 (1-u_i^2)^4}{\left(\sum_{-1<u_i<1} (1-u_i^2)(1-5u_i^2)\right) \times \max\left(1, -1 + \sum_{-1<u_i<1} (1-u_i^2)(1-5u_i^2)\right)}}$$

$$\text{where } u_i = \frac{x_i - median}{3.7 \times s};\ s = MAD/0.6745$$

$$S_T[3.7] = 3.7 \times s_{bi}[3.7] \times \sqrt{\frac{\sum_{-1<u_i<1} u_i^2 (1-u_i^2)^4}{\left(\sum_{-1<u_i<1} (1-u_i^2)(1-5u_i^2)\right) \times \max\left(1, -1 + \sum_{-1<u_i<1} (1-u_i^2)(1-5u_i^2)\right)}}$$

$$\text{where } u_i = \frac{x_i - T_{bi}}{3.7 \times s_{bi}[3.7]}$$

(The ST formula above is from the Kafadar reference on the following page.[2])

In this example, the values are:

$t_{n-1}^{0.025} = 2.0932$

$s_{bi}\,[205.6] = 0.27043$

$S_T[3.7] = 0.04816$

Using these values, the limits for the robust reference interval are:

Lower Limit = $9.6244 - 2.0932 \bullet [(0.27043)^2 + (0.04816)^2]^{1/2}$
= $9.6244 - 2.0932 \bullet 0.2747$
= 9.05

Upper Limit = $9.6244 + 2.0932 \bullet [(0.27043)^2 + (0.04816)^2]^{1/2}$
= $9.6244 + 2.0932 \bullet 0.2747$
= 10.20

Appendix B. (Continued)

As described in Horn and Pesce,[3] with skewed populations especially, the upper limit may be better estimated by making the data symmetric, and the lower limit by Box-Cox transformation. These enhancements were applied to obtain the values in Table 7.

For more details, please consult Horn and Pesce.[3]

References for Appendix B

1 Horn PS. A biweight prediction interval for random samples. *J Am Stat Assoc.* 1988;83:249-256.

2 Kafadar K. A biweight approach to the one-sample problem. *J Am Stat Assoc.* 1982;77:416-424.

3 Horn PS, Pesce AJ. *Reference Intervals. A User's Guide.* Washington, DC: AACC Press; 2005.

Clinical and Laboratory Standards Institute consensus procedures include an appeals process that is described in detail in Section 8 of the Administrative Procedures. For further information, contact CLSI or visit our website at www.clsi.org.

Summary of Consensus Comments and Subcommittee Responses

C28-A2: *How to Define and Determine Reference Intervals in the Clinical Laboratory; Approved Guideline—Second Edition*

Section 7.4, Examples

1. Page 21: Formula (5) has misplaced decimals for means and standard deviations. It should read:

$$\text{calcium}: z = \frac{|\,9.80 - 9.57\,|}{\left(\frac{(.31)^2}{120} + \frac{(.29)^2}{120}\right)^{1/2}} = 5.94$$

Fortunately, the z value is not affected, because mathematically, the misplaced decimals cancel each other during simplification.

- **Formula (5) was updated in Section 9.4.1.**

Summary of Delegate Comments and Subcommittee Responses

C28-P3: *Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Proposed Guideline—Third Edition*

General

1. The nonparametric bootstrap method, also referred to as the Harrell-Davis bootstrap method, needs to be included as a method for establishing reference intervals. I would even suggest that the working group recommend this as the best means to establish a reference interval. I provide two additional sources to support this suggestion:

   "Among available methods for estimation of reference limits and their confidence intervals, the bootstrap method probably is the most reliable one." Solberg HE. Establishment and Use of Reference Values. In: Burtis CA, Ashwood ER, Bruns DE, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. 4th ed. St. Louis, MO: Elsevier Saunders; 2006:425-448.

   "In the end, we suggest either of two procedures: the nonparametric Harrell-Davis bootstrap method or a Box-Cox transform followed, if necessary, by a second transform to eliminate residual kurtosis." Harris EK, Boyd JC. *Statistical Basis of Reference Values in Laboratory Medicine*. New York, NY: Marcel Dekker; 1995.

- **The working group agrees that the Harrell-Davis bootstrap method is an excellent method for establishing reference intervals. However, as reflected in the cited references, it typically requires at least 100 reference individual samples (per partition) and thus fails to address the practical problem of how to deal with sample sizes smaller than 100. In addition, the Harrell-Davis method involves more complex computations (and demands more expertise) than the simple nonparametric method recommended in this document.**

  Two other comments are in order regarding the Harrell-Davis method. It is indeed more efficient (ie, provides potentially narrower confidence limits at comparable numbers of data points) than the simple nonparametric technique highlighted in this document. However, being a weighted sum of all observations, the Harrell-Davis estimate is more susceptible to outliers, especially when the number of points involved is fewer than 150.

2. Clarify that the robust method is a "transform-to-Gaussian" parametric procedure. While the robust method provides a new way to estimate the center and spread of the transformed distribution, it will not satisfy users who object to the parametric approach to reference interval estimation.

- **Although it uses the form of a Gaussian estimator as a template, the robust method described in this document is not a "transform-to-Gaussian" parametric procedure. As detailed in the citations, the upper limit uses no transformation at all. The robust method requires symmetry. It deals with it, for the upper limit, by flipping the sample. For the lower limit, it uses Box-Cox, which usually does a reasonable transformation for symmetry. Note that, because achieving normality is not necessary, the robust approach does not require further transformation to remove kurtosis.**

  The working group's reason for including the robust approach is to deal specifically with the common practical problem of small sample sizes. Our preference remains that the nonparametric approach be used, but it requires a minimum of 120 points.

3. Update the references and give a more general description of how modern statistical techniques have improved reference interval estimation. Include robust estimation and bootstrap methods in that discussion. Some key references that should be updated/included are:

   – the chapter on reference values in the newest edition of Tietz (cited above);
   – Harris and Boyd's book (cited above);
   – Shultz, et al. Improved reference-interval estimation. *Clin Chem*. 1985;31:1974-1978; and
   – Linnet. Nonparametric estimation of reference intervals by simple and bootstrap-based procedures. *Clin Chem*. 2000;31:867-869.

- **The suggested changes to the references were made. The discussion was modified to include explicit mention of modern statistical techniques such as the Harrell-Davis bootstrap method. For reasons mentioned in the response to comment 1, the working group continues to highlight the robust method.**

4. Modify the claims about the number of individuals needed to establish reference intervals using the robust method. The following are specific statements and examples I find objectionable:

- **The discussion was modified substantially in accordance with this comment. All responses addressing these concerns are clearly specified below.**

  Section 1 (now Section 2, Introduction), second bullet: With the robust method, it should be "possible to establish reference intervals on far fewer than the 120 individuals required by nonparametric techniques." (Note that the nonparametric bootstrap method does not require 120 individuals; the term "far fewer" is too vague.)

- **The second bullet was completely rewritten.**

  Section 8.1 (now Section 9.1, Minimum Number of Reference Values), final paragraph: "Although there is no specific minimum number of required observations, it is recommended that the number of samples not be less than 20." (Twenty is simply too few—see the next comment.)

- **The sentence was completely rewritten as follows: "With respect to the robust method, there is no specific minimum number of required observations."**

  Section 8.4.2 (now Section 9.4.2, Calculation of Reference Intervals on Small Sample Sizes Using the Robust Method): The example with 20 data points needs to be removed. It is actually an illustration of why we should not try to estimate reference intervals with so few data points. The 90% confidence interval for the lower reference limit is not only wider than the reference interval itself, it also includes the estimated upper reference limit. This is not useful.

- **The example with 20 points was removed; it was replaced with examples using 80 points. However, for illustration purposes, Appendix B, with appropriate disclaimers, demonstrates the robust method calculations using 20 observations.**

  Section 8.4.2 (now Section 9.4.2), last paragraph: I disagree that a point estimate can be "reasonable" when its uncertainty, revealed by the width of its confidence interval, is so large.

- **The section was completely rewritten. A new emphasis is placed on the importance of the width of the confidence limits relative to the width of the reference interval.**

5. Overall, I am unconvinced that the robust method can be used with fewer than 80 values for the range of data distributions that we see with actual reference interval applications. This is based on my review of Horn, Pesce, and Copeland in Clinical Chemistry (1998), Horn and Pesce in *Reference Intervals: A User's Guide* (2005), and the examples and appendixes of the current draft of the CLSI guideline. I have many concerns, but two in particular are: the confidence intervals around the reference limits are too wide for smaller sample sizes; the medical doctors with whom I work would never accept results from as few as 20 reference values, especially if I tell them I am also screening for outliers, testing and applying transformations, and possibly using two different methods to estimate the upper and lower reference limits of the same distribution. In fact, the lack of parsimony with the robust method would cause problems regardless of the sample size.

   In my own work, I prefer a minimum of 100 subjects, but will use as few as about 80 with the bootstrap method. I have not published a formal justification for this number, but I have found it gives reasonable results and has been accepted by the medical doctors with whom I confer. I agree that the robust method example for calcium in women in the CLSI document with $n = 80$ provides an acceptably narrow set of confidence limits.

- **The working group does not advocate the use of small sample sizes. We strongly endorse the use of the simple nonparametric procedure with $n = 120$ (per partition). With highly skewed distributions, even higher n's would be recommended.**

New updates on using the robust method were added in the document to specifically address the common practical problem of having small sample sizes. Clearly, if one has 120 observations, the working group recommends using the simple nonparametric approach. For smaller sample sizes, the working group endorses the use of the robust approach; furthermore, as indicated in the document, 100 observations are better than 80, which are better than 60, and so on.

6. We would like to add a sample size calculation for establishing reference intervals that is based on requirements on the accuracy of the resulting interval (ie, limits for the probability of exceeding certain deviations between true and estimated interval).

   The following approach is **distribution free** (ie, the results are valid for any continuous distribution).

   Consider the following two **accuracy requirements** on the nonparametric reference interval, which should be fulfilled with a prescribed probability of confidence 1−α:

   A. Requirement for estimation of the 90% central interval 0.05 0.95 [x ,x]:

   Simultaneously, 0.02 0.05 0.08 $x \leq x^{\wedge} \leq x$ and 0.92 0.95 0.98 $x \leq x^{\wedge} \leq x$ (the true percentages below the sample percentiles are within nominal percentages ±3% absolute).

   B. Requirement for estimation of the 95% central interval 0.025 0.975 [x ,x]:

   Simultaneously, 0.01 0.025 0.05 $x \leq x^{\wedge} \leq x$ and 0.95 0.975 0.99 $x \leq x^{\wedge} \leq x$ (the true percentages below the sample percentiles are within 1% to 5% at the low end, 95% to 99% at the high end).

   Then, the sample sizes according to Table 1 are required:

Table 1. Required Sample Size N		
Level of Confidence, 1 – α (%)	**Requirement**	
	A	**B**
80	122	164
90	186	241
95	261	311

   Of course, it is possible to extend this table in all directions, eg, at level of confidence 90%, N = 70 is sufficient to fulfill the following requirement C for the interval 0.05 0.95 [x, x], which is weaker than requirement A above:

   C. Weakened version of requirement A:

   Simultaneously, 0.01 0.05 0.10 $x \leq x^{\wedge} \leq x$ and 0.90 0.95 0.99 $x \leq x^{\wedge} \leq x$ (the true percentages below the sample percentiles are within 1% to 10% at the low end, 90% to 99% at the high end).

- **This comment provides interesting insight into an important aspect of this guideline. As noted in Section 9.1, 120 is the minimum number of observations needed to obtain 90% confidence intervals (CIs) on the lower and upper reference limits with the simple nonparametric method. With larger numbers, one can obtain 95% and 99% CIs for the reference limits. However, these calculations make no statement as to the percentages of the population that fall within a specific range of those limits.**

  In contrast, according to example B in the comment, with 241 samples, one could ensure, with 90% certainty, that the lower limit (the 2.5th percentile) would lie in the interval from the 1st percentile to the 5th percentile of the reference population. Note that this approach does not guarantee the width of the concentration range of the CI. Validation of the numbers cited in the tables for this comment is beyond the scope of this guideline.

7. Limitations of more robust parametric approach: We think the suggestion of the draft guideline, that a parametric approach, based on outlier-robust estimators for standard deviation (SD) and the center of the distribution will enable a laboratory or manufacturer to establish reference intervals on the basis of 20, 40, or 80 measurements, is misleading. The most severe limitation of the parametric approach is not the outlier sensitivity of the underlying estimators of SD and center of distribution, but rather the unknown relation between SD and the distance between percentiles like 0.95 *x* and the center of the distribution, this relation being dependent on the shape of the distribution.

   Hence, if the distribution is non-normal, the more robust estimators are inconsistent to the same degree as the nonrobust original parametric estimators. Identifying a correct normalizing pretransformation (as suggested in a short remark in the draft guideline) with sufficient confidence requires quite a large sample size for itself (not just 20 or 40).

   It is not to be questioned that the more robust parametric approach is useful, if the underlying distribution is known to be a normal distribution that is contaminated with a certain fraction of randomly occurring outliers, but the suggested small sample sizes seem inappropriate even in this case for reference intervals.

- **The robust method described in this document does not require that the underlying distribution be normal and therefore, does not require a "normalizing pretransformation." However, the working group acknowledges the main thrust of this comment: that the numbers of reference samples needed to achieve a CI of a given width (in the original measurement units) around a reference limit is a function of the skewness of the distribution. The more highly skewed the population, the higher the number of reference samples needed to establish CIs of a given width for the reference limit in the longer tail of the distribution.**

8. Calculation of minimal sample size should include considerations with respect to the precision of the obtained reference interval.

   Example: Requiring that both of the obtained limits (upper and lower limit) of the sample-based reference interval are within the accuracy limits,

   True percentile ± c,
   c = 0.1*(length of the true reference interval)

   with a prescribed confidence probability 1-alpha, requires the following sample sizes for the central 90%-reference range (x_05 to x_95):

1-alpha (%)	N
90%	163
95%	215
99%	345

   Here, true percentile and length of the true reference interval refer to the true population percentiles and the corresponding interval. The above computations are for the normal distribution.

- **As noted in the response to comment 6, the simple nonparametric method, with *n* = 120, provides 90% CIs for the reference limits, but it does not give any assurance about the widths of those CIs or what percentage of the population they encompass. The thrust of this comment is that one may want to control the width of the CI around the reference limit (whereas, in comment 6, the idea was to ensure the percentiles). In this comment, the commenter proposes that the size of the 90% CI (± c) around the reference limit should not exceed some fraction c of the whole reference interval (upper limit minus lower limit). First proposed by Linnet (*Clin Chem*. 1987;33:381-386), this criterion for sample size determination is fine and easily computed as long as one knows the underlying form of the reference distribution (eg, normal or log normal), but it is difficult to use when the underlying reference distribution is unknown.**

In this case, the working group (using the approach on page 69 of the Harris/Boyd reference) for c = 0.1 (which gives a ratio of the width of the CI to the width of the reference interval of 0.2) finds different numbers from those cited above. For a central 90% reference interval for a reference population with an underlying normal distribution as proposed by the author, the working group obtains the following table of estimates:

1-alpha (%)	N
90%	59
95%	84
99%	118

9. The title should probably emphasize that this is for quantitative laboratory tests so those working with qualitative tests can look elsewhere for guidance.

- **The working group believes the title is appropriate for this document. In addition, the tagline on the title page states, "this document contains guidelines for determining reference values and reference intervals for quantitative clinical laboratory tests."**

10. I am concerned about the frequent references to "eliminating" outliers. While the authors state in the fourth paragraph of Section 8.2 (now Section 9.2), "Unless outliers are known to be aberrant observations …, the emphasis should be on retaining rather than deleting them," they make repeated comments about deleting outliers throughout the rest of the document (leaving the impression that deleting outliers should be standard practice). Even with sample sizes as small as 20, comments about repeatedly testing and deleting outliers are emphasized. The results of deleting these outliers is to create a tighter reference range, which would likely result in more patients with false-positive results. If the next tier of testing is not very invasive, there may be no negative consequences to having the tighter limits. If, however, the next tier of testing is invasive, then the wider limits may be more appropriate. Also, with small sample sizes, those "outliers" are more likely to be expressions of the extremes of normal. More emphasis on retaining outliers should be made throughout this document.

- **As noted in the comment, the working group believes (and states) that the emphasis should be on retaining data points. The reason so much attention is devoted to the methods of dealing with outliers is to ensure that practical, well-defined, and validated methods are used in their evaluation. The working group believes this level of detail is needed in this document, as opposed to simply providing a reference to other documents, to help ensure there is no confusion as to the proper way to proceed.**

Section 2, Introduction

11. Rewrite the second bullet. Refer to both bootstrap methods and robust estimation as modern statistical techniques that have improved reference interval estimation. Be clear that the robust method is an enhanced parametric method, ultimately relying on finding a near-Gaussian transformation of the data.

- **The second bullet of the Introduction was completely rewritten. The robust method is not "an enhanced parametric technique" and does not rely on "finding a near-Gaussian transformation of the data."**

Section 6.1, New Analyte or Analytical Method (formerly Section 5.1); Section 6.2, Multicenter Reference Interval Studies (formerly Section 5.2); Section 7.4, Selection of Reference Individuals (formerly Section 6.4); and Section 7.4.1, Direct Sampling Techniques (formerly Section 6.4.1)

12. Section 5.1, last two paragraphs; Section 5.2, bullets one and four; Section 6.4, first paragraph, last sentence and third paragraph; and Section 6.4.1, first paragraph, first sentence and throughout the second paragraph: *a priori* and *a posteriori* appear at these various locations, which is okay, but it also seems logical to include these two terms in the Definitions section for user friendliness.

- **The definitions of *a priori* and *a posteriori* were added to Section 4.2, Definitions.**

Section 7.1, Introduction (formerly Section 6.1)

13. First paragraph, second sentence: A more accurate description is needed.

    Replace "underlined terms" with "italicized terms": Section 3 of this document gives definitions of the above ~~underlined~~ italicized terms.

- **The sentence was updated as suggested. "Underlined terms" was changed to "italicized terms."**

Section 7.3, Sample Questionnaire (formerly Section 6.3)

14. Figure 1: Number each item for user friendliness and especially for ease in doing data analysis. Demographics may be exempt from numbering.

- **Each item in Figure 1 was numbered as suggested.**

Section 8, Preanalytical and Analytical Considerations (formerly Section 7)

15. Third paragraph, fourth sentence: Change in verbiage to maintain sentence structure: "Subjects using pharmacologic agents causing enzyme induction should…"

- **The sentence was updated as suggested.**

Section 8.1, Subject Preparation (formerly Section 7.1)

16. Fourth paragraph: "Refer to reference 3." Initially, this phrase is not clear. It is also appears inconsistent within text referencing of CLSI documents (eg, see CLSI in text reference in the first paragraph on page 14, Section 7.2.1). The recommended change given here in response to this comment is consistent with other similar citations in this document. Change "refer to reference 3" to refer to the published document of Solberg and PetitClerc, reference 3 at the end of this document.

- **Reference of Solberg and PetitClerc's publication was clarified.**

Section 8.2, Specimen Type, Collection, Handling, and Storage (formerly Section 7.2)

17. Importance of specimen integrity must also be considered. Add the following as sentence 3 of the second paragraph: "Specimen integrity must also be considered." (This should be inserted right before the sentence, "Fluids should be clear…"

- **The statement, "Specimen integrity must also be considered" was added to the paragraph.**

Section 8.3, Analytical Method Characteristics (formerly Section 7.3)

18. First paragraph, second sentence: Change in verbiage to maintain sentence structure "…is critical. The methods used must be…interference characteristics, and especially its..."

- **The change was made as suggested.**

Section 9, Analysis of Reference Values (formerly Section 8)

19. Include a description of the bootstrap method in Section 8. Include the references listed above. I can provide draft text if needed.

- **See the response to comment 1.**

Section 9.1, Minimum Number of Reference Values (formerly Section 8.1)

20. Rework the material in Section 8.1. The sample size requirements of the simple nonparametric method can be described with fewer words. While it can be pointed out that parametric methods like the robust method can have very small minimum requirements (we need a certain number of observations to estimate transformation, location, and spread parameters), there are in fact limits based on the acceptable uncertainty of the reference limit estimates. Similarly, the bootstrap method would yield unhelpfully wide confidence intervals with small numbers, especially for heavy-tailed data.

- **See the response to comment 1.**

21. First paragraph, first sentence: Change in verbiage to maintain sentence structure "In use of the robust method…"

- **The working group believes the current wording is appropriate.**

Section 9.2, Treatment of Outlying Observations (formerly Section 8.2)

22. Eighth paragraph, first sentence: Change in verbiage to maintain sentence structure "…the one-third rule (or any similar D/R rule) may fail to label…"

- **The statement was updated as suggested.**

Section 9.4, Examples (formerly Section 8.4)

23. In this section, calculate confidence intervals along with the reference limit estimates. I suggest calculating the reference intervals using the same data for all methods described (simple nonparametric, bootstrap nonparametric, robust parametric). I would be happy to provide bootstrap estimates and confidence intervals, if needed.

- **See the response to comment 1.**

  The working group chose not to provide details of the nonparametric bootstrap method not because it is flawed, but because it does not offer much of an advantage in practical terms over the two methods highlighted in the document.

Section 9.4.2, Calculation of Reference Intervals on Small Sample Sizes Using the Robust Method (formerly Section 8.4.2)

24. I do not believe the material in Section 8.4.2 belongs in the body of the guideline. This is promoting a specific methodology more than it is helping users judge the adequacy of their data, including sample size, to estimate reference intervals.

- **See the responses to comments 1, 4, and 5.**

Section 9.5, Confidence Intervals for Reference Limits (formerly Section 8.5)

25. Move Section 8.5 so it comes before the examples. I suggest a stronger emphasis on looking at the width of the confidence intervals. If they are unacceptably wide, there are not enough data to estimate reference intervals given the observed reference distribution.

- **The working group believes the current document structure is appropriate. See the responses to comments 1, 6, and 8.**

Section 10.1, Transference: Comparability of the Analytical System (formerly Section 9.1)

26. For transference regression in Section 9.1, suggested criteria for acceptable fit should be added to include:

a. checking assumptions about the appropriateness of the linear regression;
b. explicitly stating acceptability of regression ($r^2 = 0.90$); and
c. cautions about the use of Excel for linear regression analysis (see McCullough, Bruce [Foresight, 2006]).

- **The working group shares the concerns about the use of linear regression in connection with transference of reference intervals, which is why the document includes two specific examples as well as three caveats. With respect to the specific recommendations in this comment:**

  a) The suggestion regarding "checking assumptions about the appropriateness of the linear regression" seems to be covered in the current three caveats.
  b) It is difficult to state and justify an explicit level of r^2 for acceptability; this would depend on the specific circumstances of each case.
  c) The examples used do not explicitly recommend the use of Excel, and the third caveat specifically states that "linear regression may not always be the best, or most appropriate, tool to use."

  As a result of these concerns, at the conclusion of the section on transference is the statement that the working group "strongly encourages laboratories to validate the reference interval with a small sample (eg, *n* = 20)."

Section 11.2, Validation: Using Small Numbers of Reference Individuals (formerly Section 10.2)

27. The graph background in Figure 4 is gray and the simulated lines are colored. These do not print well when using a black and white printer. (I may not be alone in not having ready access to a color printer or sufficient monitor access to view this document online.) Would it be possible to make the graph background no color or white and make the different lines more distinguishable?

- **Figure 4 of the document was updated to make the background color white for better readability.**

28. Sixth paragraph, first sentence: Change in verbiage to maintain sentence structure: "The Mann-Whitney U (M-W U) test is very powerful..."

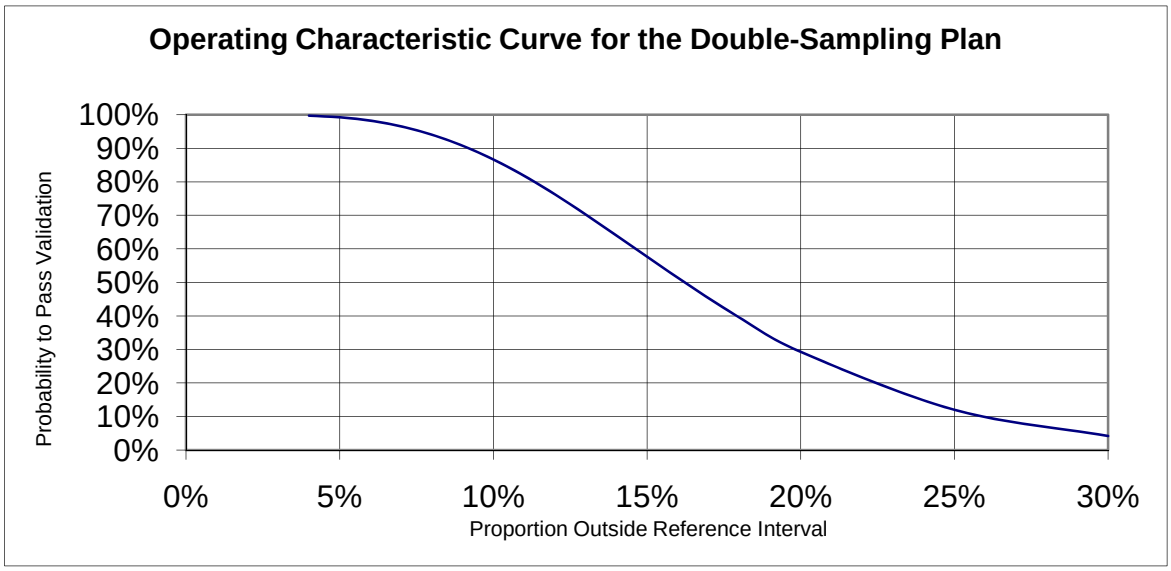
- **The statement was updated as suggested.**

29. Seventh paragraph, first sentence: Change in verbiage to maintain sentence structure: "The Siegel-Tukey (S-T) test for differences in scale is a very powerful method…"

- **The statement was updated as suggested.**

30. A different approach to analyzing the performance of the proposed validation rule in Section 11.2 is to determine the probability of passing the validation based on the proportion of the user's population outside the reference interval, as shown in the accompanying graph. This should replace the current discussion and graphs in Section 11.2.

    Recall that the rule is to accept the proposed reference interval if no more than 2 of 20 samples are outside the proposed interval; if three or four values are outside the interval, then analyze an additional 20 samples, and again accept the proposed interval if no more than 2 of these 20 samples are outside the proposed interval.



- **This is a very helpful comment. It supplements and complements the current discussion in which the analysis focuses on how sensitive different statistical tests are at detecting various perturbations of the data. In contrast, this diagram explains the performance of the proposed validation rule as a function of the proportion of the population outside the reference interval. In other words, if one's own population has 20%, rather than 5%, of individuals outside the proposed reference interval, how likely is it that one would (inappropriately) accept the proposed reference interval? According to the graph, the probability would be just under 30% (rather than just over 99%).**

  The working group has performed an initial validation of the figures in this graph and strongly endorses the concept. This type of analysis warrants more extensive validation and perhaps inclusion in the next revision of this guideline. It also could be expanded by modeling different rules, as well.

Appendix B. Robust Calculation

31. This would be a good and useful document for every clinical laboratory, as well as diagnostic companies. I found a few mistakes in the document. On page 43, there is a table that is based on the data of Table 7. If you look at columns 3 and 4, you will see the same replicating data in these two columns. It may have happened just because of pasting the same formula to the column in your Excel sheet or using previously used T_{bi} instead of the next new one for column 4 in your formula. It seems mostly probable because of the first, since the T_{bi} underweight iteration column 4 is different than column 3. There are a few more corrections needed; I corrected the table on this page and highlighted the changes in yellow.

Xi		Wi1	Wi2	Wi3	Wi4	Wi5	Wi6
8.9		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9.2		0.2193	0.1798	0.1671	0.1630	0.1616	0.1612
9.4		0.7518	0.7132	0.6998	0.6953	0.6938	0.6933
9.4		0.7518	0.7132	0.6998	0.6953	0.6938	0.6933
9.5		0.9346	0.9121	0.9038	0.9010	0.9001	0.8998
9.5		0.9346	0.9121	0.9038	0.9010	0.9001	0.8998
9.5		0.9346	0.9121	0.9038	0.9010	0.9001	0.8998
9.6		1.0000	0.9982	0.9968	0.9963	0.9961	0.9960
9.6		1.0000	0.9982	0.9968	0.9963	0.9961	0.9960
9.6		1.0000	0.9982	0.9968	0.9963	0.9961	0.9960
9.6		1.0000	0.9982	0.9968	0.9963	0.9961	0.9960
9.7		0.9346	0.9540	0.9598	0.9616	0.9622	0.9624
9.7		0.9346	0.9540	0.9598	0.9616	0.9622	0.9624
9.7		0.9346	0.9540	0.9598	0.9616	0.9622	0.9624
9.7		0.9346	0.9540	0.9598	0.9616	0.9622	0.9624
9.7		0.9346	0.9540	0.9598	0.9616	0.9622	0.9624
9.8		0.7518	0.7883	0.8001	0.8039	0.8051	0.8056
9.9		0.4913	0.5366	0.5517	0.5567	0.5583	0.5589
9.9		0.4913	0.5366	0.5517	0.5567	0.5583	0.5589
10.2		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Tbi	9.6	9.6163	9.6218	9.6236	9.6242	9.6244	9.6244

- **The table was updated with the correct values.**

32. I have a question about the formula on page 44 that I also sent in JPEG format. What is the meaning of (1,-1+ and remaining), especially the numbers that are underlined and red colored (as you also find it in the red-lined area in the JPEG file)?

were done in Excel.)

$$s_{bi}[205.6] = 205.6 \times s \times \sqrt{\frac{n \sum_{-1<u_i<1} u_i^2 (1-u_i^2)^4}{\left(\sum_{-1<u_i<1} (1-u_i^2)(1-5u_i^2)\right) \times \max\left(1, -1 + \sum_{-1<u_i<1} (1-u_i^2)(1-5u_i^2)\right)}}$$

- **An explanatory first paragraph was added to the text in Appendix B.**

NOTES

The Quality Management System Approach

Clinical and Laboratory Standards Institute (CLSI) subscribes to a quality management system approach in the development of standards and guidelines, which facilitates project management; defines a document structure via a template; and provides a process to identify needed documents. The approach is based on the model presented in the most current edition of CLSI document HS01—*A Quality Management System Model for Health Care.* The quality management system approach applies a core set of "quality system essentials" (QSEs), basic to any organization, to all operations in any health care service's path of workflow (ie, operational aspects that define how a particular product or service is provided). The QSEs provide the framework for delivery of any type of product or service, serving as a manager's guide. The QSEs are:

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

C28-A3c addresses the QSEs indicated by an "X." For a description of the other documents listed in the grid, please refer to the Related CLSI Reference Materials section on the following page.

Documents & Records	Organization	Personnel	Equipment	Purchasing & Inventory	Process Control	Information Management	Occurrence Management	Assessments—External & Internal	Process Improvement	Customer Service	Facilities & Safety
H11		H11	H11	H03	X C24 C49 EP09 GP16 H03 H04 H11 H18 H21 M29 X05				H11		H03 H11 M29

Adapted from CLSI document HS01—*A Quality Management System Model for Health Care*.

Path of Workflow

A path of workflow is the description of the necessary steps to deliver the particular product or service that the organization or entity provides. For example, CLSI document GP26—*Application of a Quality Management System Model for Laboratory Services* defines a clinical laboratory path of workflow, which consists of three sequential processes: preexamination, examination, and postexamination. All clinical laboratories follow these processes to deliver the laboratory's services, namely quality laboratory information.

C28-A3c addresses the clinical laboratory path of workflow steps indicated by an "X." For a description of the other documents listed in the grid, please refer to the Related CLSI Reference Materials section on the following page.

Preexamination				**Examination**			**Postexamination**	
Examination ordering	Sample collection	Sample transport	Sample receipt/processing	Examination	Results review and follow-up	Interpretation	Results reporting and archiving	Sample management
H03 H11	GP16 H03 H04 H11 H21	GP16 H03 H11 H18 H21	GP16 H03 H11 H18	H03 H18	H03			

Adapted from CLSI document HS01—*A Quality Management System Model for Health Care*.

Related CLSI Reference Materials*

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

C49-A **Analysis of Body Fluids in Clinical Chemistry; Approved Guideline (2007).** This document provides guidance for the application of widely available measurement procedures for testing body fluids and for reporting and interpreting those results. It emphasizes defining the common clinical situations for this use; acceptable practice for measuring analytes without extended method verification for abnormal body fluid; influence of biologic and analytic variation on interpretation of results; and variability in comparing results between different instrument manufacturers. This document does not consider serum, plasma, whole blood, or fluids for which assays typically have performance claims in the measurement procedure documentation.

EP09-A2 **Method Comparison and Bias Estimation Using Patient Samples; Approved Guideline—Second Edition (2002).** This document addresses procedures for determining the bias between two clinical methods, and the design of a method comparison experiment using split patient samples and data analysis.

GP16-A2 **Urinalysis and Collection, Transportation, and Preservation of Urine Specimens; Approved Guideline—Second Edition (2001).** This document addresses procedures for testing urine, including materials and equipment; macroscopic/physical evaluation; chemical analysis; and microscopic analysis. In addition, a step-by-step outline for collecting, transporting, and storing specimens is included.

H03-A6 **Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard—Sixth Edition (2007).** This document provides procedures for the collection of diagnostic specimens by venipuncture, including line draws, blood culture collection, and venipuncture in children.

H04-A6 **Procedures and Devices for the Collection of Diagnostic Capillary Blood Specimens; Approved Standard—Sixth Edition (2008).** This document provides a technique for the collection of diagnostic capillary blood specimens, including recommendations for collection sites and specimen handling and identification. Specifications for disposable devices used to collect, process, and transfer diagnostic capillary blood specimens are also included.

H11-A4 **Procedures for the Collection of Arterial Blood Specimens; Approved Standard—Fourth Edition (2004).** This document provides principles for collecting, handling, and transporting arterial blood specimens to assist with reducing collection hazards and ensuring the integrity of the arterial specimen.

H18-A3 **Procedures for the Handling and Processing of Blood Specimens; Approved Guideline—Third Edition (2004).** This document includes criteria for preparing an optimal serum or plasma sample and for the devices used to process blood specimens.

H21-A5 **Collection, Transport, and Processing of Blood Specimens for Testing Plasma-Based Coagulation Assays and Molecular Hemostasis Assays; Approved Guideline—Fifth Edition (2008).** This document provides procedures for collecting, transporting, and storing blood; processing blood specimens; storage of plasma for coagulation testing; and general recommendations for performing the tests.

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

X05-R **Metrological Traceability and Its Implementation; A Report (2006).** This document provides guidance to manufacturers for establishing and reporting metrological traceability.

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

NOTES

NOTES

Active Membership (as of 1 October 2008)

Sustaining Members

Abbott
American Association for Clinical Chemistry
AstraZeneca Pharmaceuticals
Bayer Corporation
BD
Beckman Coulter, Inc.
bioMérieux, Inc.
CLMA
College of American Pathologists
GlaxoSmithKline
Ortho-Clinical Diagnostics, Inc.
Pfizer Inc
Roche Diagnostics, Inc.

Professional Members

American Academy of Family Physicians
American Association for Clinical Chemistry
American Association for Laboratory Accreditation
American Association for Respiratory Care
American Medical Technologists
American Society for Clinical Laboratory Science
American Society for Microbiology
American Type Culture Collection
ASCP
Association of Public Health Laboratories
Associazione Microbiologi Clinici Italiani (AMCLI)
Canadian Society for Medical Laboratory Science
COLA
College of American Pathologists
College of Medical Laboratory Technologists of Ontario
College of Physicians and Surgeons of Saskatchewan
Elkin Simson Consulting Services
ESCMID
Family Health International
Hong Kong Accreditation Service Innovation and Technology Commission
International Federation of Biomedical Laboratory Science
International Federation of Clinical Chemistry
Italian Society of Clinical Biochemistry and Clinical Molecular Biology
JCCLS
The Joint Commission
Malaria Research Training Center
National Society for Histotechnology, Inc.
Ontario Medical Association Quality Management Program-Laboratory Service
RCPA Quality Assurance Programs PTY Limited
Serbian Society of Microbiology
SIMeL
Sociedad Espanola de Bioquimica Clinica y Patologia Molecular
Sociedade Brasileira de Analises Clinicas
Sociedade Brasileira de Patologia Clinica
Turkish Society of Microbiology
USP
World Health Organization

Government Members

BC Centre for Disease Control
CAREC
Centers for Disease Control and Prevention
Centers for Disease Control and Prevention, China
Centers for Disease Control and Prevention – Ethiopia
Centers for Disease Control and Prevention – Namibia
Centers for Disease Control and Prevention – Nigeria
Centers for Disease Control and Prevention – Tanzania
Centers for Medicare & Medicaid Services
Centers for Medicare & Medicaid Services/CLIA Program
Chinese Committee for Clinical Laboratory Standards
Department of Veterans Affairs
DFS/CLIA Certification
FDA Center for Biologics Evaluation and Research
FDA Center for Devices and Radiological Health
FDA Center for Veterinary Medicine
Health Canada
Indiana State Department of Health
Institute of Tropical Medicine Dept. of Clinical Sciences
Joanneum Research Forschungsgesellschaft mbH
Kern County Department of Health Services
Meuhedet Central Lab
Ministry of Health and Social Welfare – Tanzania
Namibia Institute of Pathology
National Cancer Institute
National Center of Infectious and Parasitic Diseases (Bulgaria)
National Health Laboratory Service (South Africa)
National Institute of Standards and Technology
National Pathology Accreditation Advisory Council (Australia)
New York City Office of Medical Examiner
New York State Department of Health
NJ State Department of Health and Senior Services
Ontario Ministry of Health
Orange County Health Care Agency - Public Health Lab
Pennsylvania Dept. of Health
Saskatchewan Health-Provincial Laboratory
Scientific Institute of Public Health
State of Alabama
State of Wyoming Public Health Laboratory
University of Iowa, Hygienic Lab
US NAMRU-3
US Naval Medical Research Unit #3
USAMC - AFRIMS

Industry Members

3M Medical Division
AB Biodisk
Abbott
Abbott Diabetes Care
Abbott Molecular Inc.
Abbott Point of Care Inc.
Access Genetics
Advanced Liquid Logic
Affymetrix, Inc. (W. Sacramento, CA)
Ammirati Regulatory Consulting
Anapharm, Inc.
Anna Longwell, PC
Arpida Ltd.
A/S Rosco
Associated Regional & University Pathologists
Astellas Pharma
AstraZeneca Pharmaceuticals
Aviir, Inc.
Axis-Shield PoC AS
Bayer Corporation – West Haven, CT
Bayer HealthCare, LLC, Diagnostic Div. – Elkhart, IN
BD
BD Diagnostic Systems
BD Vacutainer Systems
Beckman Coulter, Inc.
Beth Goldstein Consultant (PA)
BG Medicine Inc
Bioanalyse, Ltd.
Bio-Development S.r.l.
Biohit Oyj.
Biomedia Laboratories SDN BHD
bioMérieux, Inc. (MO)
BioMérieux, Inc. (NC)
BioMérieux Spain
Bio-Rad Laboratories, Inc. – France
Bio-Rad Laboratories, Inc. – Irvine, CA
Bio-Reference Laboratories
Blaine Healthcare Associates, Inc.
Canon U.S. Life Sciences, Inc.
Cempra Pharmaceuticals, Inc.
Central States Research Centre, Inc.
Cepheid
Chen & Chen, LLC (IQUUM)
The Clinical Microbiology Institute
Comprehensive Cytometric Consulting
Copan Diagnostics Inc.
Cosmetic Ingredient Review
Cubist Pharmaceuticals
Cumbre Inc.
Dade Behring Marburg GmbH – A Siemens Company
Dahl-Chase Pathology Associates PA
David G. Rhoads Associates, Inc.
DiagnoCure US, GP
Diagnostica Stago
Docro, Inc.
Dynacare Laboratory
Dynacare NW, Inc - Seattle
Eiken Chemical Company, Ltd.
Elanco Animal Health
Emisphere Technologies, Inc.
Eurofins Medinet
Fio
Focus Diagnostics
Gen-Probe
GeneNews
Genomic Health, Inc.
Genzyme Diagnostics
GlaxoSmithKline
GR Micro LTD
Greiner Bio-One Inc.
Habig Regulatory Consulting
Helena Laboratories
HistoGenex N.V.
Icon Laboratories, Inc.
Ikonisys Inc.
Immunicon Corporation
Instrumentation Laboratory (MA)
Instrumentation Laboratory (NY)
Japan Assn. of Clinical Reagents Industries
Johnson & Johnson Pharmaceutical Research and Development, L.L.C.
Kaiser Permanente
K.C.J. Enterprises
Krouwer Consulting
LabNow, Inc.
Laboratory Corporation of America
Laboratory Specialists, Inc.
LifeLabs
LifeScan, Inc. (a Johnson & Johnson Company)
LipoScience
Maine Standards Company, LLC
Medical Device Consultants, Inc.
Merck & Company, Inc.
Merial Limited
Metabolon Inc.
Micromyx, LLC
Monogen, Inc.
Nanosphere, Inc.
Nihon Koden Corporation
Nissui Pharmaceutical Co., Ltd.
NJK & Associates, Inc.
NorDx – Scarborough Campus
NovaBiotics (Aberdeen, UK)
Novartis Institutes for Biomedical Research
Nucryst Pharmaceuticals
Olympus America, Inc.
OneWorld-Lab, LLC
Opti Scan Bio Medical Assoc.
Optimer Pharmaceuticals, Inc.
OralDNA Labs, Inc.
Oregon Translational Research and Drug Development Institute
Ortho-Clinical Diagnostics, Inc. (Rochester, NY)
Ortho-McNeil, Inc.
Paratek Pharmaceuticals, Inc.
PathCare
PerkinElmer Genetics, Inc
Pfizer Animal Health
Pfizer Inc
Pfizer Italia Srl
Phadia AB
Powers Consulting Services
PPD
QML Pathology
QSE Consulting
Qualtek Clinical Laboratories
Quest Diagnostics, Incorporated
Quintiles Laboratories, Ltd.
Radiometer America, Inc.
Reametrix Inc.
Replidyne
Roche Diagnostics GmbH
Roche Diagnostics, Inc.
Roche Diagnostics (M) Sdn. Bhd
Roche Diagnostics Shanghai Ltd.
Roche Molecular Systems
Sanofi Pasteur
Sarstedt, Inc.
Schering Corporation
Sequenom, Inc.
Siemens Healthcare Diagnostics
Siemens Medical Solutions Diagnostics (CA)
Siemens Medical Solutions Diagnostics (DE)
Siemens Medical Solutions Diagnostics (NY)
Specialty Ranbaxy Ltd
Sphere Medical Holding Limited
Stirling Medical Innovations
Streck Laboratories, Inc.
Sysmex America, Inc. (Mundelein, IL)
Sysmex Corporation (Japan)
Targanta Therapeutics, Inc
TheraDoc
Theravance Inc.
Third Wave Technologies, Inc.
ThromboVision, Inc.
Transasia Bio-Medicals Limited
Trek Diagnostic Systems
Upside Endeavors, LLC
Ventana Medical Systems Inc.
Vital Diagnostics S.r.l.
Watin-Biolife Diagnostics and Medicals
Watson Pharmaceuticals
Wellstat Diagnostics, LLC
Wyeth Research
XDX, Inc.
YD Consultant

Trade Associations

AdvaMed
Japan Association of Clinical Reagents Industries (Tokyo, Japan)

Associate Active Members

3rd Medical Group
5th Medical Group/SGSL (ND)
22 MDSS (KS)
36th Medical Group/SGSL (Guam)
55th Medical Group/SGSAL (NE)
59th MDW/859th MDTS/MTL Wilford Hall Medical Center (TX)
81st MDSS/SGSAL (MS)
Academisch Ziekenhuis-VUB (Belgium)
ACL Laboratories (IL)
ACL Laboratories (WI)
Adams County Hospital (OH)
Adena Regional Medical Center Hospital (OH)
The AGA Khan University Hospital (Pakistan)
Air Force Institute for Operational Health (TX)
Akron's Children's Hospital (OH)
Al Hada Armed Forces Hospital/TAIF/KSA (Saudi Arabia)
Alameda County Medical Center (CA)
Albany Medical Center Hospital (NY)
Albemarle Hospital (NC)
All Children's Hospital (FL)
Allegheny General Hospital (PA)
Allegiance Health (MI)
Alpena General Hospital (MI)
Alta Bates Summit Medical Center (CA)
Antelope Valley Hospital District (CA)
Arkansas Children's Hospital (AR)
Arkansas Dept of Health Public Health Laboratory (AR)
Arkansas Methodist Medical Center (AR)
Asan Medical Center (Korea)
Asante Health System (OR)
Asiri Group of Hospitals Ltd. (Sri Lanka)
Aspirus Wausau Hospital (WI)
Atlantic City Medical Center (NJ)
Auburn Regional Medical Center (WA)
Aultman Hospital (OH)
Avera McKennan (SD)
Az Sint-Jan (Belgium)
Azienda Ospedale Di Lecco (Italy)
Baffin Regional Hospital (Canada)
Baptist Hospital for Women (TN)
Baptist Hospital of Miami (FL)
Baton Rouge General (LA)
Baxter Regional Medical Center (AR)
Bay Regional Medical Center (MI)
BayCare Health System (FL)
Baylor Health Care System (TX)
Bayou Pathology, APMC (LA)
Baystate Medical Center (MA)
B.B.A.G. Ve U. AS., Duzen Laboratories (Turkey)
Beebe Medical Center (DE)
Belfast HSS Trust Royal Victoria Hospital (Ireland)
Beloit Memorial Hospital (WI)
Ben Taub General Hospital (TX)
The Bermuda Hospitals Board (Bermuda)
Beth Israel Medical Center (NY)
Bonnyville Health Center (Canada)
Boston Medical Center (MA)
Boulder Community Hospital (CO)
Brantford General Hospital (Canada)
Bridgeport Hospital (CT)
Broward General Medical Center (FL)
Cadham Provincial Laboratory-MB Health (Canada)
Calgary Health Region (Canada)
California Pacific Medical Center (CA)
Cambridge Health Alliance (MA)

Camden Clark Memorial Hospital (WV)
Canadian Science Center for Human and Animal Health (Canada)
CANNON AFB (NM)
Cape Cod Hospital (MA)
Cape Fear Valley Medical Center Laboratory (NC)
Capital Coast Health (New Zealand)
Capital Health - Regional Laboratory Services (Canada)
Capital Health/QE II Health Sciences Centre (Nova Scotia, Canada)
Capital Health System Mercer Campus (NJ)
Carilion Labs Charlotte (NC)
Carl R. Darnall Army Medical Center (TX)
Carolinas Healthcare System (NC)
Cavan General Hospital (Ireland)
Cedars-Sinai Medical Center (CA)
Central Kansas Medical Center (KS)
Central Texas Veterans Health Care System (TX)
Centralized Laboratory Services (NY)
Centre Hospitalier Anna-Laberge (Canada)
Centre Hospitalier Brome-Missisquoi-Perkins (Canada)
Chaleur Regional Hospital (Canada)
Chang Gung Memorial Hospital (Taiwan)
Changhua Christian Hospital (Taiwan)
Charleston Area Medical Center (WV)
The Charlotte Hungerford Hospital (CT)
Chatham - Kent Health Alliance (Canada)
Chesapeake General Hospital (VA)
Chester County Hospital (PA)
Children's Healthcare of Atlanta (GA)
Children's Hospital and Regional Medical Center (WA)
Children's Hospital Medical Center (OH)
Children's Hospital of Philadelphia (PA)
Children's Hospitals and Clinics (MN)
Children's Medical Center (OH)
Children's Medical Center (TX)
Children's Memorial Hospital (IL)
The Children's Mercy Hospital (MO)
Childrens Hosp. – Kings Daughters (VA)
Childrens Hospital Los Angeles (CA)
Childrens Hospital of Wisconsin (WI)
Chilton Memorial Hospital (NJ)
Christiana Care Health Services (DE)
Christus St. John Hospital (TX)
CHU Sainte-Justine (Quebec, Canada)
City of Hope National Medical Center (CA)
Clarian Health – Clarian Pathology Laboratory (IN)
Clinical Labs of Hawaii (HI)
Clinton Memorial Hospital (OH)
College of Physicians and Surgeons of Alberta (Canada)
Columbia Regional Hospital (MO)
Commonwealth of Virginia (DCLS) (VA)
The Community Hospital (OH)
Community Hospital of the Monterey Peninsula (CA)
Community Medical Center (NJ)
Community Memorial Hospital (WI)
Consultants Laboratory of WI LLC (WI)
Contra Costa Regional Medical Center (CA)
Cook Children's Medical Center (TX)
Cornwall Community Hospital (Canada)
Covance CLS (IN)
The Credit Valley Hospital (Canada)
Creighton University Medical Center (NE)
Crozer-Chester Medical Center (PA)
Cumberland Medical Center (TN)
Darwin Library NT Territory Health Services (Australia)
David Grant Medical Center (CA)
Daviess Community Hospital (IN)
Deaconess Hospital Laboratory (IN)
Dean Medical Center (WI)
DeWitt Healthcare Network (USA Meddac) (VA)
DHHS NC State Lab of Public Health (NC)
Diagnostic Laboratory Services, Inc. (HI)
Diagnostic Services of Manitoba (Canada)
Diagnósticos da América S/A (Brazil)
Diaz Gill-Medicina Laboratorial S.A. (Paraguay)
Dimensions Healthcare System Prince George's Hospital Center (MD)
DMC University Laboratories (MI)
Dr. Erfan & Bagedo General Hospital (Saudi Arabia)
DRAKE Center (OH)
Driscoll Children's Hospital (TX)
DUHS Clinical Laboratories Franklin Site (NC)
Dundy County Hospital (NE)
Dunn Memorial Hospital (IN)
Durham VA Medical Center (NC)
DVA Laboratory Services (FL)
E. A. Conway Medical Center (LA)
East Central Health (Canada)
East Georgia Regional Medical Center (GA)
Eastern Health Pathology (Australia)
Easton Hospital (PA)
Edward Hospital (IL)
Effingham Hospital (GA)
Eliza Coffee Memorial Hospital (AL)
Emory University Hospital (GA)
Evangelical Community Hospital (PA)
Evans Army Community Hospital (CO)
Evanston Hospital (IL)
Exeter Hospital (NH)
Federal Medical Center (MN)
Fletcher Allen Health Care (VT)
Fleury S.A. (Brazil)
Florida Hospital (FL)
Florida Hospital Waterman (FL)
Fort St. John General Hospital (Canada)
Forum Health Northside Medical Center (OH)
Fox Chase Cancer Center (PA)
Frankford Hospital (PA)
Fraser Health Authority Royal Columbian Hospital Site (Canada)
Fresenius Medical Care/Spectra East (NJ)
Fundacio Joan Costa Roma Consorci Sanitari de Terrassa (Spain)
Garden City Hospital (MI)
Garfield Medical Center (CA)
Geisinger Medical Center (Danville, PA)
Genesis Healthcare System (OH)
George Washington University Hospital (DC)
Georgetown University Hospital - Medstar Health (MD)
Ghent University Hospital (Belgium)
Good Samaritan Hospital (OH)
Good Shepherd Medical Center (TX)
Grana S.A. (TX)
Grand Strand Reg. Medical Center (SC)
Gundersen Lutheran Medical Center (WI)
Guthrie Clinic Laboratories (PA)
Haga Teaching Hospital (Netherlands)
Hagerstown Medical Laboratory (MD)
Halton Healthcare Services (Canada)
Hamad Medical Corporation (Qatar)
Hamilton Regional Laboratory Medicine Program (Canada)
Hanover General Hospital (PA)
Harford Memorial Hospital (MD)
Harris Methodist Fort Worth (TX)
Hartford Hospital (CT)
Health Network Lab (PA)
Health Partners Laboratories Bon Secours Richmond (VA)
Health Sciences Research Institute (Japan)
Health Waikato (New Zealand)
Heidelberg Army Hospital (APO, AE)
Helen Hayes Hospital (NY)
Hema-Quebec (Canada)
Hennepin Faculty Association (MN)
Henry Ford Hospital (MI)
Henry M. Jackson Foundation (MD)
Henry Medical Center, Inc. (GA)
Hi-Desert Medical Center (CA)
Highlands Medical Center (AL)
Hoag Memorial Hospital Presbyterian (CA)
Holy Cross Hospital (MD)
Holy Family Medical Center (WI)
Holy Name Hospital (NJ)
Holy Spirit Hospital (PA)
Hopital Cite de La Sante de Laval (Canada)
Hopital du Haut-Richelieu (Canada)
Hôpital Maisonneuve - Rosemont (Montreal, Canada)
Hôpital Sacré-Coeur de Montreal (Quebec, Canada)
Hopital Santa Cabrini Ospedale (Canada)
Hospital das Clinicas-FMUSP (Brazil)
Hospital Dirino Espirito Santa (Portugal)
The Hospital for Sick Children (Canada)
Hôtel Dieu Grace Hospital Library (Windsor, ON, Canada)
Hotel-Dieu De Levis (Canada)
Imelda Hospital (Belgium)
Indian River Memorial Hospital (FL)
Inova Fairfax Hospital (VA)
Institut fur Stand. und Dok. im Med. Lab. (Germany)
Institut National de Santé Publique du Quebec Centre de Doc. – INSPQ (Canada)
Institute of Clinical Pathology and Medical Research (Australia)
Integrated Regional Laboratories South Florida (FL)
International Health Management Associates, Inc. (IL)
IWK Health Centre (Canada)
Jackson County Memorial Hospital (OK)
Jackson Health System (FL)
Jackson Purchase Medical Center (KY)
Jacobi Medical Center (NY)
John C. Lincoln Hospital (AZ)
John H. Stroger, Jr. Hospital of Cook County (IL)
John T. Mather Memorial Hospital (NY)
Johns Hopkins Medical Institutions (MD)
Johns Hopkins University (MD)
Johnson City Medical Center Hospital (TN)
Kadlec Medical Center (WA)
Kaiser Permanente (CA)
Kaiser Permanente (MD)
Kaiser Permanente (OH)
Kaiser Permanente Medical Care (CA)
Kantonsspital Aarau AG (Switzerland)
Kenora-Rainy River Reg. Lab. Program (Canada)
King Abdulaziz University Hospital (Saudi Arabia)
King Fahad Medical City (Saudi Arabia)
King Fahad National Guard Hospital KAMC – NGHA (Saudi Arabia)
King Faisal Specialist Hospital (MD)
King Hussein Cancer Center (Jordan)
Kings County Hospital Center (NY)
Kingston General Hospital (Canada)
Lab Medico Santa Luzia LTDA (Brazil)
Labette Health (KS)
Laboratory Alliance of Central New York (NY)
LabPlus Auckland Healthcare Services Limited (New Zealand)
Labway Clinical Laboratory Ltd (China)
Lafayette General Medical Center (LA)
Lakeland Regional Laboratories (MI)
Lakeland Regional Medical Center (FL)
Landspitali Univ. Hospital (Iceland)
Landstuhl Regional Medical Center (APO, AE)
Langley Air Force Base (VA)
LeBonheur Children's Medical Center (TN)
Legacy Laboratory Services (OR)
Lethbridge Regional Hospital (Canada)
Lewis-Gale Medical Center (VA)
Licking Memorial Hospital (OH)
LifeBridge Health Sinai Hospital (MD)
LifeLabs Medical Laboratory Services (Canada)
Long Beach Memorial Medical Center (CA)
Los Angeles County Public Health Lab (CA)
Louisiana Office of Public Health Laboratory (LA)
Louisiana State University Medical Ctr. (LA)
Maccabi Medical Care and Health Fund (Isreal)
Mackay Memorial Hospital (Taiwan)
Madison Parish Hospital (LA)
Mafraq Hospital (UAE)
Magnolia Regional Health Center (MS)
Makerere University Walter Reed Project Makerere University Medical School (Uganda)
Maricopa Integrated Health System (AZ)
Marquette General Hospital (MI)
Marshfield Clinic (WI)
Martha Jefferson Hospital (VA)
Martin Luther King, Jr. Harbor Hospital (CA)
Martin Memorial Health Systems (FL)
Mary Hitchcock Memorial Hospital (NH)
Mary Imogene Bassett Hospital (NY)
Massachusetts General Hospital (MA)
Mayo Clinic (MN)
Mayo Clinic Florida (FL)
Mayo Clinic Scottsdale (AZ)
Meadows Regional Medical Center (GA)
Medecin Microbiologiste (Canada)
Medical Center Hospital (TX)
Medical Center of Louisiana at NO-Charity (LA)
Medical Center of McKinney (TX)
Medical Centre Ljubljana (Slovenia)
Medical College of Georgia - Biomedical and Radiological Technologies (GA)
Medical College of Virginia Hospital (VA)
Medical Specialists (IN)
Medical Univ. of South Carolina (SC)
MediCorp - Mary Washington Hospital (VA)
Memorial Hermann Healthcare System (TX)
Memorial Hospital at Gulfport (MS)
Memorial Medical Center (IL)
Memorial Medical Center (PA)
Memorial Regional Hospital (FL)
Mercy Hospital (ME)
Methodist Dallas Medical Center (TX)
Methodist Hospital (Houston, TX)
Methodist Hospital (San Antonio, TX)
Methodist Hospital Pathology (NE)
MetroHealth Medical Center (OH)
Metropolitan Hospital Center (NY)
Metropolitan Medical Laboratory, PLC (IA)
The Michener Inst. for Applied Health Sciences (Canada)
Mid Michigan Medical Center – Midland (MI)
Middelheim General Hospital (Belgium)
Mike O'Callaghan Federal Hospital (NV)
Mississippi Baptist Medical Center (MS)
Mississippi Public Health Lab (MS)
Monongalia General Hospital (WV)
Montreal General Hospital (Quebec, Canada)
Mt. Sinai Hospital - New York (NY)
MuirLab (CA)
National Cancer Center (S. Korea)
National Healthcare Group (Singapore)
National Institutes of Health, Clinical Center (MD)
National Naval Medical Center (MD)
National University Hospital Department of Laboratory Medicine (Singapore)
Nationwide Children's Hospital (OH)
Naval Hospital Great Lakes (IL)
Naval Hospital Oak Harbor (WA)
Naval Medical Center Portsmouth (VA)
NB Department of Health (Canada)
The Nebraska Medical Center (NE)
New England Baptist Hospital (MA)
New England Fertility Institute (CT)
New Lexington Clinic (KY)
New York Presbyterian Hospital (NY)
New York University Medical Center (NY)
Newark Beth Israel Medical Center (NJ)
Newton Memorial Hospital (NJ)
North Carolina Baptist Hospital (NC)
North Coast Clinical Laboratory, Inc. (OH)
North District Hospital (Hong Kong, China)
North Mississippi Medical Center (MS)
North Shore Hospital Laboratory (New Zealand)
North Shore-Long Island Jewish Health System Laboratories (NY)
Northridge Hospital Medical Center (CA)
Northside Hospital (GA)
Northwest Texas Hospital (TX)
Northwestern Memorial Hospital (IL)
Norton Healthcare (KY)
Ochsner Clinic Foundation (LA)
Ohio State University Hospitals (OH)
Onze Lieve Vrouw Ziekenhuis (Belgium)
Ordre Professionel des Technologistes Medicaux du Quebec (Quebec, Canada)
Orebro University Hospital (Sweden)
Orlando Regional Healthcare System (FL)
Ospedale Casa Sollievo Della Sofferenza – IRCCS (Italy)
The Ottawa Hospital (Canada)
Our Lady of Lourdes Medical Center (NJ)
Our Lady of Lourdes Reg. Medical Ctr. (LA)
Our Lady's Hospital for Sick Children (Ireland)
Palmetto Health Baptist Laboratory (SC)
Pathlab (IA)
Pathology and Cytology Laboratories, Inc. (KY)
Pathology Associates Medical Laboratories (WA)
Pathology Associates of Boone (NC)
Penn State Hershey Medical Center (PA)
Pennsylvania Hospital (PA)
The Permanente Medical Group (CA)
Peterborough Regional Health Centre (Canada)
Piedmont Hospital (GA)
Pitt County Memorial Hospital (NC)
Prairie Lakes Hospital (SD)
Presbyterian Hospital – Laboratory (NC)
Presbyterian Hospital of Dallas (TX)

Presbyterian/St. Luke's Medical Center (CO)
Prince County Hospital (Canada)
Providence Alaska Medical Center (AK)
Providence Health Care (Canada)
Provincial Health Services Authority (Vancouver, BC, Canada)
Provincial Laboratory for Public Health (Edmonton, AB, Canada)
Queen Elizabeth Hospital (Canada)
Queen Elizabeth Hospital (China)
Queensland Health Pathology Services (Australia)
Quest Diagnostics, Inc
Quest Diagnostics Laboratories (WA)
Quincy Hospital (MA)
Rady Children's Hospital San Diego (CA)
Redington-Fairview General Hospital (ME)
Regional Health Authority Four (RHA4) (Canada)
Regions Hospital (MN)
Reid Hospital & Health Care Services (IN)
Renown Regional Medical Center (NV)
Research Medical Center (MO)
Riverside Regional Medical Center (CA)
Riyadh Armed Forces Hospital, Sulaymainia (Saudi Arabia)
Robert Wood Johnson University Hospital (NJ)
Rockford Memorial Hospital (IL)
Roxborough Memorial Hospital (PA)
Royal Victoria Hospital (Canada)
Rush North Shore Medical Center (IL)
SAAD Specialist Hospital (Saudi Arabia)
Sacred Heart Hospital (FL)
Sacred Heart Hospital (WI)
Sahlgrenska Universitetssjukhuset (Sweden)
Saint Francis Hospital & Medical Center (CT)
Saint Mary's Regional Medical Center (NV)
Saints Memorial Medical Center (MA)
St. Anthony Hospital (OK)
St. Anthony's Hospital (FL)
St. Barnabas Medical Center (NJ)
St. Elizabeth Community Hospital (CA)
St. Eustache Hospital (Canada)
St. Francis Hospital (SC)
St. John Hospital and Medical Center (MI)
St. John's Hospital (IL)
St. John's Hospital & Health Ctr. (CA)
St. John's Mercy Medical Center (MO)
St. John's Regional Health Center (MO)
St. Joseph Health Center (MO)
St. Joseph Mercy Hospital (MI)
St. Joseph's Medical Center (CA)
St. Joseph's Regional Medical Center (NJ)
St. Jude Children's Research Hospital (TN)
St. Louis University Hospital (MO)
St. Luke's Hospital (IA)
St. Luke's Hospital (PA)
St. Martha's Regional Hospital (Canada)
St. Mary Medical Center (CA)
St. Mary's Hospital (WI)
St. Tammany Parish Hospital (LA)
Sampson Regional Medical Center (NC)
Samsung Medical Center (Korea)
San Francisco General Hospital-University of California San Francisco (CA)
San Juan Regional Medical Group (NM)
Sanford USP Medical Center (SD)
Santa Clara Valley Medical Center (CA)
Saudi Aramco Medical (Saudi Arabia)
Schneck Medical Center (IN)
Scott & White Memorial Hospital (TX)
Scott Air Force Base (IL)
Seoul National University Hospital (Korea)
Seton Medical Center (CA)
Sheik Kalifa Medical City (UAE)
Shore Memorial Hospital (NJ)
Shriners Hospitals for Children (SC)
Singapore General Hospital (Singapore)
SJRMC Plymouth Laboratory (IN)
Sky Lakes Medical Center (OR)
South Bend Medical Foundation (IN)
South Miami Hospital (FL)
Southern Health Care Network (Australia)
Southern Maine Medical Center (ME)
Speare Memorial Hospital (NH)
Spectrum Health - Blodgett Campus (MI)
Stanford Hospital and Clinics (CA)
State of Connecticut Department of Public Health (CT)
State of Hawaii Department of Health (HI)
State of Washington-Public Health Labs (WA)
Stillwater Medical Center (OK)
Stony Brook University Hospital (NY)
Sudbury Regional Hospital (Canada)
Suncoast Medical Clinic (FL)
Sunnybrook Health Science Center (ON, Canada)
Sunrise Hospital and Medical Center (NV)
Sydney South West Pathology Service Liverpool Hospital (Australia)
T.J. Samson Community Hospital (KY)
Taipei Veterans General Hospital (Taiwan)
Taiwan Society of Laboratory Medicine (Taiwan)
Tallaght Hospital (Ireland)
Tartu University Clinics (Estonia)
Texas Children's Hospital (TX)
Texas Department of State Health Services (TX)
Thomason Hospital (TX)
Timmins and District Hospital (Canada)
The Toledo Hospital (OH)
Touro Infirmary (LA)
Tri-Cities Laboratory (WA)
Trident Medical Center (SC)
Trinity Medical Center (AL)
Tripler Army Medical Center (HI)
Tufts New England Medical Center Hospital (MA)
Tulane Medical Center Hospital & Clinic (LA)
Turku University Central Hospital (Finland)
UC Davis Health System (CA)
UCI Medical Center (CA)
UCLA Medical Center Clinical Laboratories (CA)
UCSD Medical Center (CA)
UCSF Medical Center China Basin (CA)
UMass Memorial Medical Center (MA)
UMC of Southern Nevada (NV)
UNC Hospitals (NC)
Union Clinical Laboratory (Taiwan)
United Christian Hospital (Hong Kong)
United Clinical Laboratories (IA)
Unity HealthCare (IA)
Universita Campus Bio-Medico (Italy)
University College Hospital (Ireland)
University Hospital Center Sherbrooke (CHUS) (Canada)
University Medical Center at Princeton (NJ)
University of Alabama Hospital Lab (AL)
University of Arkansas for Medical Sci. (AR)
University of Chicago Hospitals (IL)
University of Colorado Health Sciences Center (CO)
University of Colorado Hospital (CO)
University of Iowa Hospitals and Clinics (IA)
University of Kentucky Med. Ctr. (KY)
University of Maryland Medical System (MD)
University of Miami (FL)
University of MN Medical Center - Fairview (MN)
University of MS Medical Center (MS)
University Pittsburgh Med. Ctr. (PA)
University of So. Alabama Children's and Women's Hospital (AL)
University of Texas Health Center (TX)
The University of Texas Medical Branch (TX)
University of the Ryukyus (Japan)
University of Virginia Medical Center (VA)
University of Washington (WA)
UPMC Bedford Memorial (PA)
U.S. Naval Hospital (FPO AP)
US Naval Hospital Naples (FPO)
U.S.A. Meddac (Pathology Division) (MO)
UW Hospital (WI)
UZ-KUL Medical Center (Belgium)
VA (Asheville) Medical Center (NC)
VA (Bay Pines) Medical Center (FL)
VA (Biloxi) Medical Center (MS)
VA (Chillicothe) Medical Center (OH)
VA (Cincinnati) Medical Center (OH)
VA (Dayton) Medical Center (OH)
VA (Decatur) Medical Center (GA)
VA (Indianapolis) Medical Center (IN)
VA (Long Beach) Medical Center (CA)
VA (Miami) Medical Center (FL)
VA New Jersey Health Care System (NJ)
VA Outpatient Clinic (OH)
VA (San Diego) Medical Center (CA)
VA (Seattle) Medical Center (WA)
VA (Sheridan) Medical Center (WY)
Valley Health (VA)
Vancouver Coastal Health Regional Laboratory (BC, Canada)
Vancouver Island Health Authority (SI) (Canada)
Vanderbilt University Medical Center (TN)
Via Christi Regional Medical Center (KS)
Virga Jessezieukenhuis (Belgium)
Virtua - West Jersey Hospital (NJ)
WakeMed (NC)
Walter Reed Army Medical Center (DC)
Warren Hospital (NJ)
Washington Hospital Center (DC)
Waterbury Hospital (CT)
Waterford Regional Hospital (Ireland)
Wayne Memorial Hospital (NC)
Weirton Medical Center (WV)
West China Second University Hospital, Sichuan University (P.R. China)
West Valley Medical Center Laboratory (ID)
Western Baptist Hospital (KY)
Western Healthcare Corporation (Canada)
Wheaton Franciscan & Midwest Clinical Laboratories (WI)
Wheeling Hospital (WV)
William Beaumont Army Medical Center (TX)
William Beaumont Hospital (MI)
William Osler Health Centre (Canada)
Winchester Hospital (MA)
Winn Army Community Hospital (GA)
Wisconsin State Laboratory of Hygiene (WI)
Wishard Health Sciences (IN)
Womack Army Medical Center (NC)
Woodlawn Hospital (IN)
York Hospital (PA)

940 West Valley Road ▾ Suite 1400 ▾ Wayne, PA 19087 ▾ USA ▾ PHONE 610.688.0100
FAX 610.688.0700 ▾ E-MAIL: customerservice@clsi.org ▾ WEBSITE: www.clsi.org ▾ ISBN 1-56238-682-4