C28-A2
Vol. 20 No. 13

Replaces C28-A
Vol. 15 No. 4

How to Define and Determine Reference Intervals in the Clinical Laboratory; Approved Guideline—Second 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 NCCLS consensus process.

NCCLS...

Serving the World's Medical Science Community Through Voluntary Consensus

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 healthcare 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 healthcare issues. NCCLS is based on the principle that consensus is an effective and cost-effective way to improve patient testing and healthcare services.

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

PUBLICATIONS

An NCCLS 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 NCCLS 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 NCCLS 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 (i.e., "tentative") consensus level.

Proposed An NCCLS consensus document undergoes the first stage of review by the healthcare 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.

Tentative A tentative standard or guideline is made available for review and comment only when a recommended method has a well-defined need for a field evaluation or when a recommended protocol requires that specific data be collected. It should be reviewed to ensure its utility.

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

NCCLS 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 NCCLS's established consensus procedures. Provisions in NCCLS 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 NCCLS 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 NCCLS document. Address comments to the NCCLS Executive Offices, 940 West Valley Road, Suite 1400, Wayne, PA 19087, USA.

VOLUNTEER PARTICIPATION

Healthcare professionals in all specialties are urged to volunteer for participation in NCCLS projects. Please contact the NCCLS Executive Offices for additional information on committee participation.

How to Define and Determine Reference Intervals in the Clinical Laboratory; Approved Guideline—Second Edition

Abstract

How to Define and Determine Reference Intervals in the Clinical Laboratory; Approved Guideline—Second Edition (NCCLS document C28-A2) 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.

NCCLS. *How to Define and Determine Reference Intervals in the Clinical Laboratory; Approved Guideline—Second Edition*. NCCLS document C28-A2 (ISBN 1-56238-406-6). NCCLS, 940 West Valley Road, Suite 1400, Wayne, Pennsylvania 19087-1898, USA 2000.

C28-A2
ISBN 1-56238-406-6
ISSN 0273-3099

How to Define and Determine Reference Intervals in the Clinical Laboratory; Approved Guideline—Second Edition

Volume 20 Number 13

Edward A. Sasse, Ph.D.
Basil T. Doumas, Ph.D.
W. Gregory Miller, Ph.D.
Paul D'Orazio, Ph.D.
John H. Eckfeldt, M.D., Ph.D.
Susan A. Evans, Ph.D.
Gary A. Graham, Ph.D., DABCC
Gary L. Myers, Ph.D.
Patrick J. Parsons, Ph.D.
Noel V. Stanton, M.S.

Suggested Citation

(NCCLS. *How to Define and Determine Reference Intervals in the Clinical Laboratory; Approved Guideline—Second Edition.* NCCLS document C28-A2 [ISBN 1-56238-406-6]. NCCLS, 940 West Valley Road, Suite 1400, Wayne, Pennsylvania 19087-1898, USA 2000.)

Proposed Guideline
March 1992

Approved Guideline
June 1995

Approved Guideline—Second Edition
June 2000

ISBN 1-56238-406-6
ISSN 0273-3099

Committee Membership

Area Committee on Clinical Chemistry and Toxicology

Basil T. Doumas, Ph.D. **Chairholder**	**Medical College of Wisconsin** **Milwaukee, Wisconsin**
W. Gregory Miller, Ph.D. **Vice-Chairholder**	**Virginia Commonwealth University** **Richmond, Virginia**
Paul D'Orazio, Ph.D.	Instrumentation Laboratory Lexington, Massachusetts
John H. Eckfeldt, M.D., Ph.D.	Fairview-University Medical Center Minneapolis, Minnesota
Susan A. Evans, Ph.D.	Dade Behring Inc. Deerfield, Illinois
Gary A. Graham, Ph.D., DABCC	Ortho-Clinical Diagnostics Rochester, New York
Gary L. Myers, Ph.D.	Centers for Disease Control and Prevention Atlanta, Georgia
Patrick J. Parsons, Ph.D.	New York State Department of Health Albany, New York
Noel. V. Stanton, M.S.	WI State Laboratory of Hygiene Madison, Wisconsin

Advisors

Judith T. Barr, Sc.D.	Northeastern University Boston, Massachusetts
Stanley Bauer, M.D.	Beth Israel Medical Center New York, New York
George N. Bowers, Jr., M.D.	Hartford Hospital Hartford, Connecticut
Robert W. Burnett, Ph.D.	Hartford Hospital Hartford, Connecticut
Mary F. Burritt, Ph.D.	Mayo Clinic Rochester, Minnesota
Kevin D. Fallon, Ph.D.	Instrumentation Laboratory Lexington, Massachusetts

Advisors (Continued)

Carl C. Garber, Ph.D.	Quest Diagnostics, Incorporated Teterboro, New Jersey
Harvey W. Kaufman, M.D.	Quest Diagnostics, Incorporated Teterboro, New Jersey
Jan S. Krouwer, Ph.D.	Bayer Diagnostics Medfield, Massachusetts
Victoria M. Leitz, Ph.D.	International Biomedical Consultants Hilton Head, South Carolina
Richard R. Miller, Jr.	Dade Behring Inc. Newark, Delaware
Robert F. Moran, Ph.D., FCCM, FAIC	mvi Sciences Methuen, Massachusetts
Richard B. Passey, Ph.D.	University of Oklahoma Oklahoma City, Oklahoma
Edward A. Sasse, Ph.D.	Medical College of Wisconsin Milwaukee, Wisconsin
Richard S. Schifreen, Ph.D.	Promega Corporation Madison, Wisconsin
Bette Seamonds, Ph.D.	National Academy of Clinical Biochemistry Swarthmore, Pennsylvania
Beth Ann Wise, M.T.(ASCP), M.S.Ed. *Staff Liaison*	NCCLS Wayne, Pennsylvania
Patrice E. Polgar *Editor*	NCCLS Wayne, Pennsylvania
Donna M. Wilhelm *Assistant Editor*	NCCLS Wayne, Pennsylvania

Acknowledgements

The Area Committee on Clinical Chemistry and Toxicology extends its appreciation to Edward A. Sasse, Ph.D., Chairholder of the former Subcommittee on Reference Intervals, for his help in preparing the second edition of this approved-level guideline, especially his advice on appropriate revisions and responses to the comments.

In addition, the area committee would also like to recognize the valuable contributions of the members and advisors of the Subcommittee on Reference Intervals that developed the first approved edition of this guideline.

Edward A. Sasse, Ph.D., Chairholder
Kaiser J. Aziz, Ph.D.
Eugene K. Harris, Ph.D.
Sandy Krishnamurthy
Henry T. Lee, Jr.
Andy Ruland
Bette Seamonds, Ph.D.

Advisors

Horace F. Martin, Ph.D., M.D.
John Sherwin, M.D.
Margaret Steffes

Additionally, NCCLS gratefully acknowledges James C. Boyd, M.D., of the University of Virginia Health Sciences Center for his valuable contributions of data to Section 7.4.

Active Membership
(as of 1 April 2000)

Sustaining Members

Abbott Laboratories
American Association for Clinical Chemistry
Bayer Corporation
Beckman Coulter, Inc.
BD and Company
bioMérieux, Inc.
College of American Pathologists
Dade Behring Inc.
Nippon Becton Dickinson Co., Ltd.
Ortho-Clinical Diagnostics, Inc.
Pfizer Inc
Roche Diagnostics, Inc.

Professional Members

American Academy of Family Physicians
American Association of Blood Banks
American Association for Clinical Chemistry
American Association for Respiratory Care
American Chemical Society
American Medical Technologists
American Public Health Association
American Society for Clinical Laboratory Science
American Society of Hematology
American Society for Microbiology
American Society of Parasitologists, Inc.
American Type Culture Collection, Inc.
Asociación Española Primera de Socorros (Uruguay)
Asociacion Mexicana de Bioquimica Clinica A.C.
Assn. of Public Health Laboratories
Assoc. Micro. Clinici Italiani-A.M.C.L.I.
Australasian Association of Clinical Biochemists
British Society for Antimicrobial Chemotherapy
Canadian Society for Medical Laboratory Science—Société Canadienne de Science de Laboratoire Médical
Canadian Society of Clinical Chemists
Clinical Laboratory Management Association
College of American Pathologists
College of Medical Laboratory Technologists of Ontario
College of Physicians and Surgeons of Saskatchewan
Commission on Office Laboratory Accreditation
Fundacion Bioquimica de la Provincia (Argentina)
International Association of Medical Laboratory Technologists
International Council for Standardization in Haematology
International Federation of Clinical Chemistry
International Society for Analytical Cytology
Italian Society of Clinical Biochemistry
Japan Society of Clinical Chemistry
Japanese Association of Medical Technologists (Tokyo)
Japanese Committee for Clinical Laboratory Standards
Joint Commission on Accreditation of Healthcare Organizations
National Academy of Clinical Biochemistry
National Society for Histotechnology, Inc.
Ontario Medical Association Laboratory Proficiency Testing Program
RCPA Quality Assurance Programs PTY Limited
Sociedade Brasileira de Analises Clinicas
Sociedade Brasileira de Patologia Clinica
Sociedad Espanola de Quimica Clinica

Government Members

Armed Forces Institute of Pathology
BC Centre for Disease Control
Centers for Disease Control and Prevention
Chinese Committee for Clinical Laboratory Standards
Commonwealth of Pennsylvania Bureau of Laboratories
Department of Veterans Affairs
Deutsches Institut für Normung (DIN)
FDA Center for Devices and Radiological Health
FDA Division of Anti-Infective Drug Products
Health Care Financing Administration/CLIA Program
Health Care Financing Administration
Iowa State Hygienic Laboratory
Massachusetts Department of Public Health Laboratories
National Association of Testing Authorities – Australia
National Center of Infectious and Parasitic Diseases (Bulgaria)
National Institute of Standards and Technology
Ohio Department of Health
Oklahoma State Department of Health
Ontario Ministry of Health
Saskatchewan Health-Provincial Laboratory
Scientific Institute of Public Health; Belgium Ministry of Social Affairs, Public Health and the Environment
South African Institute for Medical Research
Swedish Institute for Infectious Disease Control
Thailand Department of Medical Sciences

Industry Members

AB Biodisk
Abbott Laboratories
Abbott Laboratories, MediSense Products
Accumetrics, Inc.
Amersham Pharmacia Biotech
Ammirati Regulatory Consulting
Asséssor
AstraZeneca
Aventis
Avocet Medical, Inc.
Bayer Corporation - Elkhart, IN
Bayer Corporation - Middletown, VA
Bayer Corporation - Tarrytown, NY

Bayer Corporation – West Haven, CT
Bayer Medical Ltd.
BD
BD Biosciences – San Jose, CA
BD Biosciences – Sparks, MD
BD Consumer Products
BD Italia S.P.A.
BD VACUTAINER Systems
Beckman Coulter, Inc.
Beckman Coulter, Inc. Primary Care Diagnostics
Beckman Coulter K.K. (Japan)
Bio-Inova Life Sciences International
Biolog, Inc.
bioMérieux, Inc.
Biometrology Consultants
Bio-Rad Laboratories, Inc.
Biotest AG
Bristol-Myers Squibb Company
Canadian Reference Laboratory Ltd.
Capital Management Consulting, Inc.
CASCO•NERL Diagnostics
Checkpoint Development Inc.
Clinical Design Group Inc.
Clinical Lab Engineering
COBE Laboratories, Inc.
Combact Diagnostic Systems Ltd.
Community Medical Center (NJ)
Control Lab (Brazil)
Copan Diagnostics Inc.
Cosmetic Ingredient Review
Cubist Pharmaceuticals
Cytometrics, Inc.
Dade Behring Inc. - Deerfield, IL
Dade Behring Inc. - Glasgow, DE
Dade Behring Inc. - Marburg, Germany
Dade Behring Inc. - Sacramento, CA
Dade Behring Inc. - San Jose, CA
DAKO A/S
Diagnostic Products Corporation
DiaSorin
Eiken Chemical Company, Ltd.
Enterprise Analysis Corporation
Fort Dodge Animal Health
Gen-Probe
Glaxo-Wellcome, Inc.
Greiner Meditech, Inc.
Health Systems Concepts, Inc.
Helena Laboratories
Home Diagnostics, Inc.
Hycor Biomedical Inc.
I-STAT Corporation
Instrumentation Laboratory
International Technidyne Corporation
Johnson City Medical Center
Kendall Sherwood-Davis & Geck
Labtest Diagnostica S.A.
LifeScan, Inc. (a Johnson & Johnson Company)
Lilly Research Laboratories
Medical Automation Systems
Medical Device Consultants, Inc.
Medical Laboratory Automation Inc.
Medtronic Perfusion Systems
Merck & Company, Inc.
mvi Sciences (MA)
Nabi
Neometrics Inc.
Nichols Institute Diagnostics (Div. of Quest Diagnostics, Inc.)
Nissui Pharmaceutical Co., Ltd.
Nippon Becton Dickinson Co., Ltd.
Norfolk Associates, Inc.
Ortho-Clinical Diagnostics, Inc. (Raritan, NJ)
Ortho-Clinical Diagnostics, Inc. (Rochester, NY)
Oxoid Inc.
Pfizer Inc
Pharmacia & Upjohn
Procter & Gamble Pharmaceuticals, Inc.
The Product Development Group
Quest Diagnostics Incorporated
Quintiles, Inc.
Radiometer America, Inc.
Radiometer Medical A/S
David G. Rhoads Associates, Inc.
Roche Diagnostics GmbH
Roche Diagnostics, Inc.
Roche Laboratories (Div. Hoffmann-La Roche Inc.)
The R.W. Johnson Pharmaceutical Research Institute
Sanofi Diagnostics Pasteur
Sarstedt, Inc.
SARL Laboratoire Carron (France)
Schering Corporation
Schleicher & Schuell, Inc.
Second Opinion
SenDx Medical, Inc.
Showa Yakuhin Kako Company, Ltd.
SmithKline Beecham, S.A.
Streck Laboratories, Inc.
Sysmex Corporation (Japan)
Sysmex Corporation (Long Grove, IL)
The Toledo Hospital (OH)
Trek Diagnostic Systems, Inc.
Vetoquinol S.A.
Vysis, Inc.
Wallac Oy
Warner-Lambert Company
Wyeth-Ayerst
Xyletech Systems, Inc.
YD Consultant
Yeongdong Pharmaceutical Corporation

Trade Associations

Association of Medical Diagnostic Manufacturers
Health Industry Manufacturers Association
Japan Association Clinical Reagents Ind. (Tokyo, Japan)
Medical Industry Association of Australia

Associate Active Members

67th CSH Wuerzburg, GE (NY)
121st General Hospital (CA)
Acadiana Medical Laboratories, LTD (LA)
Advocate Laboratories (IL)
Albany Medical Center Hospital (NY)
Allegheny General Hospital (PA)
Allegheny University of the Health Sciences (PA)
Allina Laboratories (MN)
Alton Ochsner Medical Foundation (LA)
American Medical Laboratories (VA)
Anzac House (Australia)
Asan Medical Center (Korea)
Associated Regional & University Pathologists (UT)
Aurora Consolidated Laboratories (WI)
Baystate Medical Center (MA)
Brantford General Hospital (Brantford, ON, Canada)
Brazileiro De Promocao (Brazil)
Bristol Regional Medical Center (TN)
Brookdale Hospital Medical Center (NY)
Brooke Army Medical Center (TX)
Brooks Air Force Base (TX)
Broward General Medical Center (FL)
Calgary Laboratory Services
Carilion Consolidated Laboratory (VA)

CB Healthcare Complex (Sydney, NS, Canada)
Central Kansas Medical Center
Centralized Laboratory Services (NY)
Centro Diagnostico Italiano (Milano, Italy)
Champlain Valley Physicians Hospital (NY)
Children's Hospital King's Daughters (VA)
Children's Hospital (LA)
Children's Hospital (NE)
Children's Hospital Medical Center (Akron, OH)
Children's Hospital of Philadelphia (PA)
Clendo Lab (Puerto Rico)
CLSI Laboratories (PA)
Commonwealth of Kentucky
Commonwealth of Virginia (DCLS)
CompuNet Clinical Laboratories (OH)
Consolidated Laboratory Services (CA)
Covance Central Laboratory Services (IN)
Danish Veterinary Laboratory (Copenhagen, Denmark)
Danville Regional Medical Center (VA)
Dean Medical Center (WI)
Delaware Public Health Laboratory
Department of Health & Community Services (New Brunswick, Canada)
Detroit Health Department (MI)
Diagnostic Laboratory Services, Inc. (HI)
Duke University Medical Center (NC)
Durham Regional Hospital (NC)
Duzen Laboratories (Turkey)
Dynacare Laboratories - Eastern Region (Ottawa, ON, Canada)
E.A. Conway Medical Center (LA)
Elmhurst Memorial Hospital (IL)
Elyria Memorial Hospital (OH)
Emory University Hospital (GA)
Fairfax Hospital (VA)
Fairview-University Medical Center (MN)
Foothills Hospital (Calgary, AB, Canada)
Fox Chase Cancer Center (PA)
Franklin Square Hospital Center (MD)
Fresenius Medical Care/Life Chem (NJ)
Fresno Community Hospital and Medical Center
Gambro Healthcare Laboratory (FL)
GDS Technology, Inc (IN)
Grady Memorial Hospital (GA)
Greater Southeast Community Hospital (DC)
Guthrie Clinic Laboratories (PA)
Harris Methodist Fort Worth (TX)
Harris Methodist Northwest (TX)
Hartford Hospital (CT)
Hays Pathology Laboratories, P.A. (KS)
Health Alliance Laboratory (OH)
Health Network Lab (PA)
Health Sciences Centre (Winnipeg, MB, Canada)
Heartland Health System (MO)
Hinsdale Hospital (IL)
Hoag Memorial Hospital Presbyterian (CA)
Holmes Regional Medical Center (FL)
Holy Spirit Hospital (PA)
Holzer Medical Center (OH)
Hospital for Sick Children (Toronto, ON, Canada)
Hospital Israelita Albert Einstein (Brazil)
Hotel Dieu Hospital (Windsor, ON, Canada)
Huddinge University Hospital (Sweden)
Hurley Medical Center (MI)
Indiana University
Instituto Scientifico HS. Raffaele (Italy)
International Health Management Associates, Inc. (IL)
Intermountain Health Care Laboratory Services (UT)
Jacobi Medical Center (NY)
John Peter Smith Hospital (TX)
John Randolph Hospital (VA)
Johns Hopkins Medical Institutions (MD)
Johnson City Medical Center (IN)
Kaiser Permanente (CA)
Kaiser Permanente (NC)
Kantousspital (Switzerland)
Keller Army Community Hospital (NY)
Klinicni Center (Slovenia)
LabCorp (NC)
Laboratoire de Santé Publique du Quebec (Canada)
Laboratório Fleury S/C Ltda. (Brazil)
Laboratory Corporation of America (MO)
Lakeland Regional Medical Center (FL)
Lancaster General Hospital (PA)
Langley Air Force Base (VA)
Lewis-Gale Medical Center (VA)
Libero Instituto Univ. Campus BioMedico (Italy)
LAC and USC Healthcare Network (CA)
Louisiana State University Medical Center
Lutheran Hospital (WI)
Martin Luther King/Drew Medical Center (CA)
Massachusetts General Hospital (Microbiology Laboratory)
Massachusetts General Hospital (Pathology Laboratory)
Mayo Clinic Scottsdale (AZ)
MDS Metro Laboratory Services (Burnaby, BC, Canada)
Medical Center of Southern Indiana
Medical College of Virginia Hospital
Medicare/Medicaid Certification, State of North Carolina
Melrose-Wakefield Hospital (MA)
Memorial Hospital (CO)
Memorial Medical Center (Napoleon Ave., New Orleans, LA)
Memorial Medical Center (N. Jefferson Davis Pkwy., New Orleans, LA)
Memorial Medical Center (IL)
Mercy Health System (PA)
Mercy Hospital (NC)
Methodist Hospital (TX)
Methodist Hospital Indiana
Methodist Hospitals of Memphis (TN)
Michigan Department of Community Health
Montreal Children's Hospital (Canada)
Montreal General Hospital (Canada)
Mount Sinai Hospital (NY)
National University Hospital (Singapore)
Naval Surface Warfare Center (IN)
Nebraska Health System
New Britain General Hospital (CT)
New England Medical Center Hospital (MA)
The New York Hospital Medical Center of Queens

New York State Department of Health
NorDx (ME)
North Carolina Laboratory of Public Health
North Carolina School of Veterinary Medicine
North Mississippi Medical Center
Northridge Hospital Medical Center (CA)
Northwestern Memorial Hospital (IL)
Olin E. Teague Medical Center (TX)
O.L. Vrouwziekenhuis (Belgium)
Ordre professionnel des technologists médicaux du Québec
Ottawa General Hospital (Ottawa, ON, Canada)
Our Lady of Lourdes Hospital (NJ)
Our Lady of the Resurrection Medical Center (IL)
Pathology and Cytology Laboratories, Inc. (KY)
Pathology Associates Laboratories (CA)
The Permanente Medical Group (CA)
Pocono Hospital (PA)
Presbyterian Hospital (NC)
Presbyterian Hospital of Dallas (TX)
Providence Health System (OR)
Providence Seattle Medical Center (WA)
Queen Elizabeth Hospital (Prince Edward Island, Canada)
Queensland Health Pathology Services (Australia)
Quintiles Laboratories, Ltd. (GA)
Regions Hospital
Research Medical Center (MO)
Rex Healthcare (NC)
Riyadh Armed Forces Hospital (Saudi Arabia)
Robert F. Kennedy Medical Center (CA)
Royal Columbian Hospital (New Westminster, BC, Canada)
Saint Mary's Regional Medical Center (NV)
St. Alexius Medical Center (ND)
St. Anthony Hospital (CO)
St. Barnabas Medical Center (NJ)
St. Boniface General Hospital (Winnipeg, Canada)
St. Elizabeth Hospital (NJ)
St. John Hospital and Medical Center (MI)
St. John Regional Hospital (St. John, NB, Canada)
St. Joseph Medical Center (MD)
St. Joseph Hospital (NE)
St. Joseph Mercy – Oakland (MI)
St. Joseph's Hospital - Marshfield Clinic (WI)
St. Luke's Hospital (PA)
St. Luke's Regional Medical Center (IA)
St. Mary Medical Center (IN)
St. Mary of the Plains Hospital (TX)
Salina Regional Health Center (KS)
San Francisco General Hospital (CA)
Santa Cabrini Hospital (Montreal, PQ Canada)
Santa Clara Valley Medical Center (CA)
Seoul Nat'l University Hospital (Korea)
Shanghai Center for the Clinical Laboratory (China)
Shands Healthcare (FL)
SmithKline Beecham Clinical Laboratories (GA)
SmithKline Beecham Clinical Laboratories (WA)
South Bend Medical Foundation (IN)
Southern California Permanente Medical Group
South Western Area Pathology Service (Australia)
Speare Memorial Hospital (NH)
Speciality Laboratories, Inc. (CA)
Stanford Health Services (CA)
State of Washington Department of Health
Stormont-Vail Regional Medical Center (KS)
Sun Health-Boswell Hospital (AZ)
Sunrise Hospital and Medical Center (NV)
Sutter Health (CA)
Tampa General Hospital (FL)
Tripler Army Medical Center (HI)
Tulane Medical Center Hospital & Clinic (LA)
UCSF Medical Center (CA)
UNC Hospitals (NC)
Unilab Clinical Laboratories (CA)
University of Alberta Hospitals (Canada)
University of Chicago Hospitals (IL)
University of Florida
University Hospital (IN)
University Hospital (Gent) (Belgium)
University Hospital (London, Ontario, Canada)
The University Hospitals (OK)
University of Massachusetts Lowell
University of Medicine & Dentistry, NJ University Hospital
University of Michigan
University of the Ryukyus (Japan)
University of Virginia Medical Center
University of Washington
UPMC Bedford Memorial (PA)
USAF Medical Center (OH)
UZ-KUL Medical Center (Belgium)
VA (Dayton) Medical Center (OH)
VA (Denver) Medical Center (CO)
VA (Martinez) Medical Center (CA)
VA (San Diego) Medical Center (CA)
VA (Tuskegee) Medical Center (AL)
VA Outpatient Clinic (OH)
Vejle Hospital (Denmark)
Via Christi Regional Medical Center (KS)
Virginia Department of Health
Viridae Clinical Sciences, Inc. (Vancouver, BC, Canada)
Walter Reed Army Institute of Research (MD)
Warde Medical Laboratory (MI)
Warren Hospital (NJ)
Washoe Medical Center (NV)
Watson Clinic (FL)
Wilford Hall Medical Center (TX)
William Beaumont Hospital (MI)
Williamsburg Community Hospital (VA)
Wilson Memorial Hospital (NY)
Winchester Hospital (MA)
Winn Army Community Hospital (GA)
Wishard Memorial Hospital (IN)
Womack Army Medical Center (NC)
Yan Chai Hospital (P.R. China)
Yonsei University College of Medicine (Korea)
York Hospital (PA)
Zale Lipshy University Hospital (TX)

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 (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. We are indebted to the members of the IFCC committee and to the many other investigators who contributed to this discipline and upon whose knowledge we have 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.

Standard Precautions

Because it is often impossible to know what might be infectious, all human blood specimens are to be treated as infectious and handled according to “standard precautions.” Standard precautions are new guidelines that combine the major features of “universal precautions and body substance isolation” practices. Standard precautions cover the transmission of any pathogen and thus are more comprehensive than universal precautions which are intended to apply only to transmission of blood-borne pathogens. Standard precaution and universal precaution guidelines are available from the U.S. Centers for Disease Control and Prevention (*Guideline for Isolation Precautions in Hospitals*. Infection Control and Hospital Epidemiology. CDC. 1996; Vol 17;1:53-80.), [MMWR 1987;36(suppl 2S):2S-18S] and (MMWR 1988;37:377-382, 387-388). For specific precautions for preventing the laboratory transmission of blood-borne infection from laboratory instruments and materials; and recommendations for the management of blood-borne exposure, refer to NCCLS document M29—*Protection of Laboratory Workers from Instrument Biohazards and Infectious Disease Transmitted by Blood, Body Fluids, and Tissue*.

Key Words

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

How to Define and Determine Reference Intervals in the Clinical Laboratory; Approved Guideline—Second Edition

1 Introduction and Scope

This document provides diagnostic laboratories, diagnostic test manufacturers, and users of clinical laboratory tests with guidelines for determining reference values and reference intervals for quantitative clinical laboratory tests. It includes the methodological approaches and recommended procedures for establishing reliable reference intervals for use in clinical laboratory medicine. The recommendations contained in this document comprise a protocol for determining reference intervals that meet the minimum, mandatory requirements for adequate reliability and usefulness. There are situations that will require more guidance than these recommendations can provide. Such situations cannot be covered entirely in this brief document. However, in certain areas, the additional steps or efforts that would improve the reliability and accuracy of the reference interval determination are indicated. Because of the lack of uniformity in the data collection and in the methodology currently used for establishing reference intervals by clinical laboratory scientists and manufacturers, it is the subcommittee's hope that this document will provide a basic and uniform protocol for achieving a comparable level of reliability and a foundation for more elaborate studies.

The procedures for determining "health-associated" reference values or intervals derived from a reference sample group of persons who are in good health are the primary focus of the document. However, other types of reference values, for example, for other physiological or pathological conditions, could also be established in a similar manner. With attention to the selection of appropriate reference individuals and due consideration of preanalytical factors, the procedures outlined here can be followed for the determination of any type of reference interval. However, this document does not specifically address the process required to establish critical values or other medical decision limits, such as diagnostic cut-offs. These determinations require a different approach, in part, and are often based on the diagnostic sensitivity and specificity for a specific medical condition.

The various needs and requirements of reference value studies for different situations are also addressed, including:

- measurement of a new analyte
- measurement by a new or different analytical method of a previously known and measured analyte for which physiological data and other reference values may be available
- measurement of the same analyte by the same or comparable analytical method for which reference value studies from another laboratory or the manufacturer are available (transference).

The latter issue, which is referred to as “transference of reference values,” is complex. The validation of transference and the subsequent transfer of reference values will increasingly be an issue encountered by the clinical laboratory testing community as diagnostic test manufacturers and other laboratories provide appropriately determined reference value data. Approaches to this problem are not yet rigorous; consequently, this issue is discussed in terms of general recommendations and three acceptable approaches.

If a diagnostic laboratory, large or small, or a diagnostic test manufacturer has to establish a reference interval through a reference value study, the specific guidelines and procedures provided in this document should be followed. This document contains the minimum standards for an adequate and appropriate reference interval determination. If the facility is small and lacks the resources to conduct such an

appropriate reference interval determination, the only other acceptable substitute for use of a reference interval should be by transference of an already appropriately established reference interval for the same or comparable analytical system, according to the recommendations discussed later in the document.

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 EPTRV of the 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. 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 Use of Système International d'Unités (SI Units)

Although NCCLS 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 the International Federation of Clinical Chemistry (IFCC) for reporting results of clinical laboratory measurements. NCCLS documents also include the IUPAC/IFCC-recommended units of volume (L) and substance (molecular) concentration (mol/L) in parentheses, where appropriate.

3 Definitions[a]

3.1 IFCC/ICSH Definitions

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 an accepted universal terminology. A discussion and clarification of these terms follows (in Section 3.2).

Reference individual, *n* - 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.

Reference population, *n* - A group consisting of all the reference individuals. **NOTES:** a) 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 (e.g., a person may serve as a reference for himself or herself, or for another person); b) These "subject-specific" reference intervals are not addressed in this guideline.

Reference sample group, *n* - An adequate number of persons selected to represent the reference population.

Reference value, *n* - 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.

[a] Some of these definitions are found in NCCLS document NRSCL8—*Terminology and Definitions for Use in NCCLS Documents.* For complete definitions and detailed source information, please refer to the most current edition of that document.

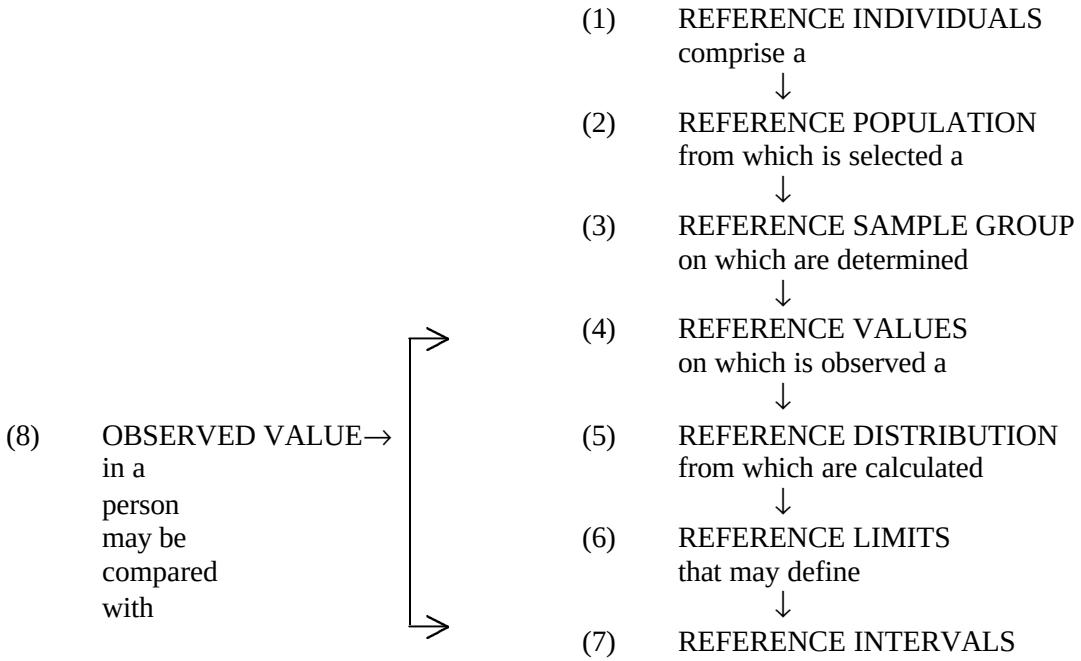
Reference distribution, *n* - The distribution of reference values. **NOTE:** Hypotheses regarding the distribution of a reference population may be tested using the reference distribution of the reference sample group and adequate statistical methods. The parameters of the hypothetical distribution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

Reference limit, *n* – A value derived from the reference distribution and used for descriptive purposes. **NOTE:** It is common practice to define a reference limit so that a stated fraction of the reference values is less than or equal, or more 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 interval, *n* - 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 (e.g., for fasting glucose the reference interval is 65 to 110 mg/dL [3.6 to 6.1 mmol/L]; in some cases, only one reference limit is important, usually an upper limit, "*x*," and the corresponding reference interval would be 0 to *x*).

Observed value, (patient laboratory test result), *n* - The value of a particular type of quantity, obtained by observation or measurement of a test subject (i.e., patient), to be compared with reference values, reference distributions, reference limits, or reference intervals.

The following scheme demonstrates the relationship between the terms defined.



3.2 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 will refer 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, 95%).

4 Protocol Outline for Obtaining Reference Values and Establishing Reference Intervals

4.1 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 or for a new analytical method for a previously measured analyte:

(1) Establish an appropriate list of biological variations and analytical interferences 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.

(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 other 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, i.e., 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 5.3) 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 canceled in the case of a discovered exclusion.

In some cases, 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, however, only after the measurements are taken rather than before.

4.2 Previously Measured Analyte

In an appropriate situation, it may be acceptable to transfer a reference interval based on a previously established, valid, reference value study from a donor laboratory or manufacturer to a receiving laboratory 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. The comparability of the analytical measuring system can be validated using the techniques discussed in NCCLS document EP9— *Method Comparison and Bias Estimation Using Patient Samples*. It may be necessary to carry out an abbreviated reference value study, as described in Section 8, to validate the transferred reference interval.

5 Selection of Reference Individuals

5.1 Introduction

This section provides guidelines and suggestions for making a reference sample group of reference individuals from a reference population.[2,7] Section 3 of this document gives definitions of the above underlined terms. Two different sampling techniques (*a priori* and *a posteriori*) will be discussed and the concepts of exclusion and partitioning will be explored. A sample questionnaire is presented.

As discussed in Section 1, it is the intent of this document to present procedures for determining "health-associated" reference values. Health is a relative condition lacking a universal definition. Defining what is to be 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 that others can evaluate the health status of that reference sample group. At a minimum, a questionnaire should be used to evaluate the health of each reference individual.

5.2 Exclusion and Partitioning

Exclusion criteria are details about the candidate reference individual which, if present, serve to keep that person from being included in the reference sample. Examples of some potential exclusion criteria may be found in Table 1 of this section. 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 and should serve to stimulate thinking about criteria necessary for the study under design. Not all reference value studies will 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	Posture when sampled
Blood group	Race
Circadian variation	Sex
Diet	Stage of menstrual cycle
Ethnic background	Stage of pregnancy
Exercise	Time of day when sampled
Fasting or nonfasting	Tobacco use
Geographic location	

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 pregnant reference sample by trimesters.

Well-designed questionnaires are one of the best ways 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 5.4.

5.3 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 subcommittee 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, some age-related changes in laboratory values also may not represent good health, e.g., cholesterol or growth hormone in the geriatric patient. Reference individuals should not be hospital or clinic patients unless absolutely necessary, as might be necessary for pediatric or geriatric values.

The terms *a priori* and *a posteriori* are used to describe two general methods of selecting reference individuals from the reference population.

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 7.1).

In *a posteriori* sampling, the process of exclusion and partitioning also takes place but in a different order (i.e., 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.

5.4 Sample Questionnaire

The questionnaire is presented in this document as an example (Figure 1). To protect the reference individuals, it is important to maintain the questionnaire information and the testing results in a confidential manner. There are several design changes that might be considered. Name, address, and phone number are included to facilitate contacting the reference individual in case the analysis uncovers some potential abnormalities. Certainly, using good medical judgment, there is an obligation to notify the person or his or her physician in such cases. 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. (The reference individual is then responsible for contacting the laboratory to determine if the testing showed any problems that require follow-up.) Certainly, the anonymous questionnaire approach is more problematic.

Another possible variation, especially in the case of a *priori* study, 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. Questionnaires, consent forms, and even the nature of the study itself may need to be reviewed by the institution's Internal Review Board or Human Subjects Committee.

6 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.[3,8]

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 (e.g., 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 5.2, and they may require separate reference intervals. In some cases, the laboratory and physician have some control over the preanalytical variables, which negates the need to separate the reference intervals. However, for specimen testing under emergency conditions, certain standardized conditions may not be applicable. Therefore, knowledge of the effects of deviation from the standardized protocol is important for appropriate interpretation of results.[9]

In general, preanalytical considerations involve two areas, namely, biological and methodological factors.[9] The biological factors include those that are of metabolic and hemodynamic origin. Procedures resulting in potential for cell damage (from physical training to venipuncture) must also be considered. Subjects using pharmacologic agents causing enzyme induction will 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). The user is referred to the IFCC checklist[9] and Tables 3 and 4 as helpful guidelines for evaluating preanalytical issues.

Measurement of the same analyte by more than one method, instrument, or system will require appropriate examinations to verify that the various methods, instruments, or systems generate comparable results. If the alternate methods or systems cannot be made to give comparable results (see Section 8.1, Transference, and NCCLS document EP9—*Method Comparison and Bias Estimation Using Patient Samples*), then separate reference intervals may have to be established, particularly if the differences in the numerical results are clinically significant.

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:_____

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

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?_____ AND WHAT?_____

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

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

DO YOU TAKE VITAMIN SUPPLEMENTS? (Y) (N)
IF YES, WHAT?_____

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

DO YOU USE TOBACCO? (Y) (N)

IF YES, WHAT FORM?_____ HOW OFTEN?_____

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

IF YES, PLEASE DESCRIBE_____

Figure 1. Sample Questionnaire

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

IF YES, WHAT FORM?_____HOW OFTEN?

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

IF YES, WHY?_____

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

IF YES, WHY?_____WHEN?_____

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

IF YES, DESCRIBE?_____

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

IF YES, WHAT?_____WHEN?_____

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

IF YES, WHAT?_____WHEN?_____

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

IF YES, WHAT?_____WHEN?_____

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

FOR WOMEN:

ARE YOU STILL MENSTRUATING? (Y) (N)

IF YES, WHEN WAS YOUR LAST PERIOD?_____

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

ARE YOU BREAST-FEEDING? (Y) (N)

ARE YOU PREGNANT? (Y) (N)

IF YES, WHAT IS YOUR DUE DATE?_____

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

Figure 1. Sample Questionnaire (Continued)

6.1 Subject Preparation

As described in Section 5, 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 which skew data. The criteria set will be dictated by the effect of biological variation on the analyte(s) of interest. Tables 3 and 4 summarize the critical factors to be considered regarding subject preparation.[3,9]

Table 3. Summary of Critical Factors

Biological Factors	Methodological Factors	Sources of Variability and Standardization
• Metabolic • Hemodynamic • Enzyme induction • Cell damage	• Specimen collection • Specimen transport • Specimen handling	• Specific factors (supine vs. upright) • Multiple factors (see Table 2)

Food ingestion before blood sampling will affect many laboratory results, either directly (changes in analyte concentration) or indirectly (lipid interference). Conversely, prolonged fasting will induce 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.[3,7,8]

Exercise and postural position during the phlebotomy procedure can change a laboratory result. The impact of postural changes is important when comparing in-patient and out-patient results and, as stated earlier, frequently will necessitate the establishment of separate reference intervals for some analytes. Other factors to consider are circadian and temporal fluctuations, which may alter analyte concentration, seasonal influences, and ethnic background.[7-9]

Many of these issues will have been eliminated by the appropriate exclusion criteria. (Refer to reference 3 for specific details about each category.)

Table 4. 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	• Transport • Clotting • Separation of serum/plasma • Storage • Preparation for analysis

6.2 Specimen Type, Collection, Handling, and Storage

The laboratory should have a manual outlining the collection, handling, and storage of specimens so that 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 NCCLS documents H3— *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture*; H4— *Procedures and Devices for the Collection of Diagnostic Blood Specimens by Skin Puncture;* H11—*Procedures for the Collection of Arterial Blood Specimens*; and H21— *Collection, Transport, and Processing of Blood Specimens for Coagulation Testing and General Performance of Coagulation Assays).*

Consideration should be given to whether the specimen should be maintained under anaerobic conditions (e.g., for ionized calcium measurements). 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. Fluids should be clear-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.

6.2.1 Blood

If blood is the specimen of choice, it will be 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 NCCLS documents H3— *Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture* and H4— *Procedures and Devices for the Collection of Diagnostic Blood Specimens by Skin Puncture).*

6.2.2 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 NCCLS document GP16— *Routine Urinalysis and Collection, Transportation, and Preservation of Urine Specimens*), 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 creatinine level.

6.2.3 Temperature

The collection and handling of some specimens may require procurement at a specific temperature (e.g., 37 ^{o}C, room temperature, or iced). In addition, preservation of some specimens (analytes) will require storage at a particular temperature or freezing, possibly at a specified temperature (-20 ^{o}C vs. -70 ^{o}C). It is essential to establish any special conditions and strictly adhere to them. In general, specimens should be processed promptly after collection. Processing will frequently entail removal of serum or plasma from the clot or red cells as quickly as possible and at a specified temperature (see NCCLS document H18— *Procedures for the Handling and Processing of Blood Specimens*).[9]

6.3 Analytical Method Characteristics

The validity of information provided by the laboratory is critical; thus, the methods chosen for specimen analysis must be described in detail, establishing their inaccuracy, imprecision, minimum detection limit, linearity, recovery, and interference characteristics.[10-12]

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 will be used. Knowledge of all the above factors will define the analytical system to be checked.

The reliability of the data generated is critical, because both the imprecision and inaccuracy of the method will determine its diagnostic utility. Therefore, also included during the determination of reference intervals is the routine use of quality control 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 NCCLS document C24—*Statistical Quality Control for Quantitative Measurements: Principles and Definitions*.) Ideally, data will be 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.[10]

7 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 have been drawn. For most analytes, the lower and upper reference limits are assumed to demarcate the estimated 2.5th and 97.5th percentiles of the underlying distribution of values, respectively. In some cases, only one reference limit is of medical importance, usually an upper limit, say the 97.5th percentile.

Two general statistical methods for determining such limits are 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 mathematical form of the probability distribution represented by 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 (i.e., "normal") probability curve. Because the reference values of many analytes do not follow the Gaussian form, use of the parametric method requires that they be transformed to some other measurement scale, which will "normalize" them. This requires selecting the most suitable transformation (e.g., log, power, or some other function of the original scale) and then testing whether, on this new scale, the reference values do indeed appear to 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 nonparametric method is far simpler, depending only on the ranks of the reference data arrayed in order of increasing size. Furthermore, the most important considerations in developing reliable reference intervals are proper selection of reference subjects, testing an adequate number of subjects, and avoidance of preanalytical sources of error, not the statistical method used to estimate the reference interval from the observed data. Therefore, the nonparametric method is recommended, although a laboratory with the required statistical and computing competence should feel free to use a parametric method if desired. Examples of the nonparametric method of estimating reference intervals for two analytes, serum calcium and alanine aminotransferase (AlaAT), are described in Section 7.4.

7.1 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 have been 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 have been 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 says, 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 (i.e., *P* = 2.5), a minimum of 39 measurements are required. The smallest observation in the sample would be the nonparametric estimate of the 2.5th percentile of the population, while the largest observation would estimate 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.[13] 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 7.5). To estimate the reference limits for these same percentiles with 95% confidence, 153 reference values are needed; for 99% confidence, 198 reference values are needed. Linnet[14] proposes that up to 700 should be obtained for highly skewed distributions of results; however, as a standard for general practice, the subcommittee supports the recommended minimum of 120 reference subjects.

This number assumes that no observations were deleted from the reference set (see Section 7.2). If aberrant or outlying observations were 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 were needed for different subclasses (by sex or age-class, for example), each such interval should be based on the recommended number (at least 120) of reference observations.

In the case of difficult-to-obtain subclass reference values for certain populations, such as newborn, pediatric, and geriatric patients, it may be difficult, if not impossible, to obtain appropriate age-related 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.

7.2 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 that all values come from the same underlying 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 almost all reference values, but that one or two arise from a different probability distribution than their fellows. 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 “aberrant” values lie outside the range of the remaining measurements, and they are easily identified as “outliers” requiring special attention.

Unless outliers are known to be aberrant observations (e.g., due to a mistake in the analysis or to a lapse in the preanalytic 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 would be only slightly changed, or not changed at all, if an extreme value were deleted.

There are many statistical techniques available for testing the atypicality of outlying observations (see Barnett and Lewis[15]). The majority of these tests 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[16] 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.[13] suggests the value 1/3 as a cut-off value; i.e., if the observed value of *D* were equal to or greater than one-third of the range *R*, the extreme observation would be deleted. For sample sizes as large as 120, this criterion is rather conservative (as Reed et al. point out). That is, it would often fail to reject outliers, which are really not part of the distribution, followed by 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 will often not be 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, we support the use of this test and cut-off 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 (i.e., all extremely large or extremely small), the one-third rule (or any similar *D/R* rule) can 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 were 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.[15]

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

7.3 Partitioning of Reference Values

The possibility that separate intervals will be 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 will be 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, will test out statistically significant if the sample sizes are large enough. Sinton et al.[17] suggest 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.

Conversely, research undertaken as part of the subcommittee's work (Harris and Boyd[18]) shows that smaller differences between subclass means can lead to situations in which the proportions of each subclass above the upper reference limit (without partitioning) or below the lower reference limit are much different than the desired 2.5% on each side. This implies the possibility of major deviations in the sensitivity and specificity for subclasses from those values obtained for the population as a whole, a condition that could seriously hamper the interpretation of laboratory results as part of the diagnostic process. (For a more in depth explanation of this issue and a general theory of reference interval determination, see Hanis EK, Body JC. *Statistical Bases of Reference Values in Laboratory Medicine.* New York: Marcel Dekker; 1995.)

Moreover, this condition can exist even when the means are identical if the standard deviations of the subclasses are in a ratio of 1.5 or more. In such a case, the wider distribution (subclass) would extend

substantially beyond the narrower distribution on both sides. However, such ratios seldom occur. In most practical cases, subclass standard deviations are approximately the same despite statistically significant differences in mean values.

After consideration of these studies, the subcommittee offers the following recommendations. First, 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. This should be done in two stages, beginning with a pilot sample of (approximately) 60 subjects in each subclass. For two subclasses (e.g., men and women or two age groups), the statistical significance of the difference between subclass means should be tested 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.[18] Assuming at least 60 subjects in each subclass, the *z*-test is essentially a nonparametric test and may be applied to the original data whether or not the values conform to a Gaussian distribution. However, 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.

The calculated statistic *z* should be compared with a "critical" value[18]

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

In addition, the larger standard deviation, for example s_2, should be checked to see whether it exceeds $1.5s_1$, or, equivalently, whether $s_2/(s_2 - s_1)$ is less than 3.[18]

For example, suppose that at the end of the first stage of sampling, the average number of reference values in each subclass is 60. Then, if the calculated *z* exceeds $z^*=3(60/120)^{½}=2.12$, or if the larger standard deviation exceeds 1.5 times the smaller standard deviation, sampling should be continued to obtain at least 120 subjects in each subclass. The *z*-test and standard deviation comparisons should be repeated. If the average number of subjects in each subclass is now 120, $z^* = 3$. If the average number in each subclass exceeds 120, then the critical value of the test statistic *z* will be greater than $z^* = 3$, as indicated by the general formula for z^* given above. For example, if 500 values were obtained in each subclass, the critical value will be $z^* = 6.12$.

At this point, if the calculated *z* value exceeds z^*, or if the larger standard deviation exceeds 1.5 times the smaller, regardless of the *z* value, then separate reference intervals should be calculated for each subclass, assuming that the difference between the two reference intervals is likely to be of importance in medical practice. If these conditions do not hold, then a single reference interval for the combined group of reference subjects should be calculated for general use.

When more than two subclasses are compared, the problem is more complicated, and the aid of a statistical consultant should be sought. An example of this situation appears in Harris et al.[19] The

following advice is offered: For three or more subclasses, the recommended statistical analysis of results is analysis of variance (ANOVA). However, a statistically significant difference among all subclass means taken together may, in fact, be due to a difference between only two of the means or between one subclass and all the others combined. Therefore, a significant *F*-test in the ANOVA should be followed by simultaneous comparison of paired means, using, for example, the Tukey test, which controls all such tests at 0.05 probability level while maintaining a high probability of detecting a real difference. The general ANOVA (allowing for unequal subclass numbers), followed by the Tukey (or other) test for paired means is available in many commonly used statistical program packages (e.g., SAS). Any pair of means whose difference is statistically significant should then be retested using the more stringent *z** criterion defined above.

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 8).

7.4 Examples

The histograms depicted in Figures 2 and 3 represent ordered values of calcium and alanine aminotransferase (AlaAT), respectively, measured in serum samples from medical students at the University of Virginia during 1987 and 1988. The data themselves are listed in rank order in Tables 5 and 6, comprising 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-like distributions; whereas those for AlaAT show considerable skewness to the right. The extreme value of 65 U/L of AlaAT (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 AlaAT values are nearly Gaussian in shape. Results for both analytes appear generally higher in men than in women, and a statistical test of separate reference intervals by sex would be of interest.

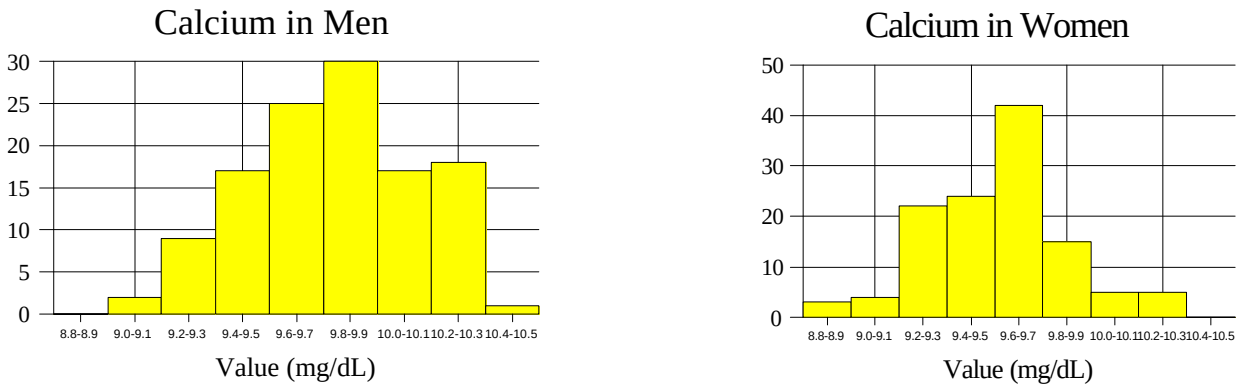


Figure 2. Calcium Histograms

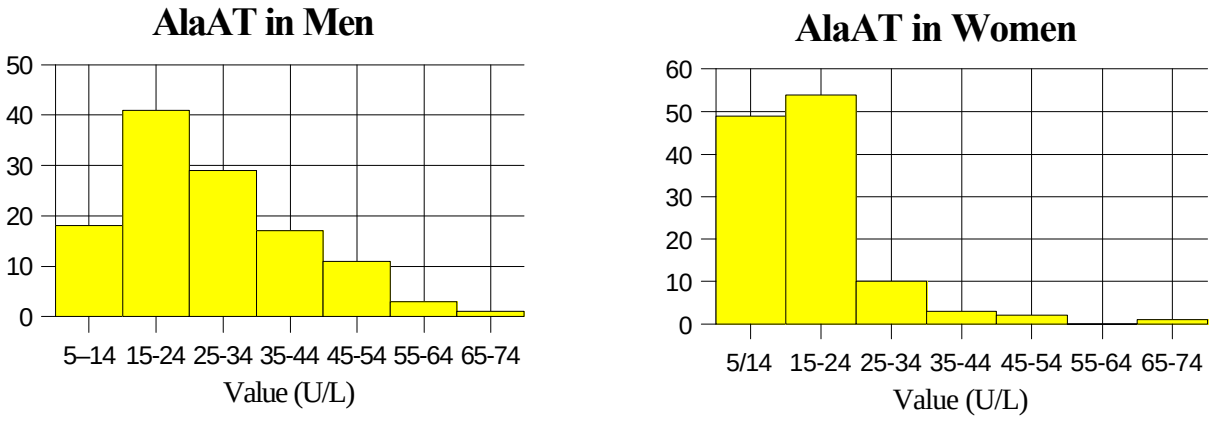


Figure 3. Alanine Aminotransferase Histograms

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

Analyte	Value (mg/dL)[a]	Frequency		
		Women	Men	Combined
Calcium	8.8	1	0	1
	8.9	2[b]	0	2
	9.0	1	0	1
	9.1	3	2	5[b]
	9.2	11	1[b]	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[c]	10
	10.1	2	10	12
	10.2	3[c]	11	14
	10.3	2	7	9[c]
	10.4	0	1	1
	10.5	0	0	0
	10.6	0	1	1
	Total	120	120	240

[a]mg/dL • 0.02495 = mmol/L.
[b]r_1 = rank value #3 (2.5th percentile), n = 120; rank value #6, n = 240.
[c]r_2 = rank value #118 (97.5th percentile), n = 120; rank value #235, n = 240.

Table 6. Frequency Distributions of Alanine Aminotransferase in 240 Medical Students by Sex

Analyte	Value (U/L)	Frequency		
		Women	Men	Combined
AlaAT	5	1	0	1
	6	3[a]	0	3
	7	1	0	1
	8	5	0	5[a]
	9	2	1	3
	10	2	2[a]	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[b]	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[b]
	55	0	2[b]	2
	62	0	1	1
	65	1	0	1
	69	0	1	1
	Total	120	120	240

[a] r_1 = rank value #3 (2.5th percentile), n = 120; rank value #6, n = 240.
[b] r_2 = rank value #118 (97.5th percentile), n = 120; rank value #235, n = 240.

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 (i.e., ordered by size). 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 $r = 0.025\ (n + 1)$, and the upper reference limit, r_2 (the 97.5th percentile), as the observation corresponding to $r = 0.975\ (n + 1)$. Since the values of r_1 and r_2 will usually not be 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 will be rounded off to these numbers:

$$r_1 = 0.025\ (121) = 3.025 \approx 3 \quad (3)$$

$$r_2 = 0.975\ (121) = 117.975 \approx 118 \quad (4)$$

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, we obtain the following 95% reference intervals:

	Calcium
Women:	8.9 to 10.2 mg/dL (2.22 to 2.54 mmol/L)
Men:	9.2 to 10.3 mg/dL (2.30 to 2.57 mmol/L)
Combined:	9.1 to 10.3 mg/dL (2.27 to 2.57 mmol/L)
	AlaAT
Women:	6 to 46 U/L
Men:	10 to 55 U/L
Combined:	8 to 54 U/L

To test the statistical significance of the differences between the mean values for calcium and AlaAT in men and women of this age group, we need the statistics given in Table 7.

Table 7. Means and Standard Deviations of Calcium and $\log_e$AlaAT in Young Men and Women, 120 in Each Subclass

Analyte	Means		Standard Deviations	
	Men	Women	Women	Men
Calcium (mg/dL)	9.8	9.57	3.1	2.9
$\log_e$AlaAT (lo U/L)*	3.2	2.78	0.46	0.44

*See Section 7.3, Partitioning of Reference Values, about the need for log transformation.

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

$$\text{calcium}: z = \frac{|98.0 - 95.7|}{\left(\frac{(3.1)^2}{120} + \frac{(2.9)^2}{120}\right)^{1/2}} = 5.94 \quad (5)$$

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

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

7.5 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 would probably yield somewhat different reference limits. A useful way of recognizing and assessing the variability in sample estimates is by computing a confidence interval for the population percentile being estimated, using information provided by the sample. In the present case, a confidence interval is a range of values that includes the true percentile (e.g., 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 confidence intervals 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 will be, in fact, healthy persons from whom reference specimens will be secured under the recommended preanalytical conditions. The reference individuals are at least randomly obtained from a described pool, e.g., laboratory employees. Therefore, the basic assumptions for the validity of confidence intervals are that the observations are obtained independently of each other and that the reference sample is representative of the population even if not drawn strictly at random.

Nevertheless, confidence intervals are useful for two reasons. First, they remind the investigator of the variability of estimates and provide a quantitative measure of this variability. Second, confidence intervals 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.

Nonparametric confidence intervals are given by the observed values corresponding to certain rank numbers. Table 8[5,13] shows the rank number defining the 90% confidence interval (CI) for the 2.5th percentile based on a given sample size.

Table 8. Rank Number Defining the 90% Confidence Interval for the 2.5th Percentile. From *Clinical Chemistry*. 1971;17:275-284 (Table 3). Reprinted with permission from the American Association for Clinical Chemistry, Inc.

No. of Sample, *n*		Rank*	
From	**To**	***a***	***b***
120	131	1	7
132	159	1	8
160	187	1	9
188	189	1	10
190	216	2	10
217	246	2	11
247	251	2	12
252	276	3	12
277	307	3	13
308	310	3	14
311	338	4	14
339	366	4	15
367	369	5	15

**a*th lowest sample value = lower limit of 90% confidence interval for 2.5 percentile in target population; *b*th lowest sample value = upper limit of 90% confidence interval for 2.5 percentile in target population. To obtain ranks corresponding to a 90% confidence interval for the 97.5 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% confidence interval for the lower reference limits. To obtain the comparable rank numbers defining the 90% confidence interval 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% confidence interval for the lower reference limit, while the largest observation is the upper limit of the 90% confidence interval for the upper reference limit.

Using the rank numbers in Table 8 and the data in Tables 5 and 6, Table 9 presents the 90% confidence intervals for the lower and upper reference limits for calcium and AlaAT.

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)	8.8-9.1	10.3-10.6
AlaAT (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 • 0.02495 = mmol/L.

8 Transference and Validation

8.1 Transference

Because the determination of reliable reference intervals can be a major and costly task, it would be very useful to be able to transfer a reference interval from one laboratory to another by some process of validation that is less costly and more convenient. As more new tests and methods are introduced in more laboratories, it is unrealistic to expect each laboratory, large or 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 can be a complex issue and requires that certain conditions be fulfilled in order to be acceptable. The requirements for acceptable transference are different for different scenarios:

(1) Transference of a reference interval for an analyte using the same (identical) type of analytical system (method and instrument):

(a) Within the same laboratory
(b) From one laboratory to another

- For the same geographic region and same demographic population of test subjects
- For a different geographic region or different demographic population of test subjects

(2) Transference of a reference interval for an analyte measured by a different analytical system (different method or different instrument):

(a) Within the same laboratory
(b) From one laboratory to another

- For the same geographic region and same demographic population of test subjects
- For a different geographic region or different demographic population of test subjects.

Assuming that the original reference value study was done properly, the transference of the respective reference interval involves two primary and distinct problems: the comparability of the analytical system and the comparability of the test subject population.

If an appropriately established reference interval for an analyte already exists for the population of subjects being tested using the clinical laboratory's current system, then transference of the reference interval within the same laboratory to an alternate method/instrument becomes a question of the comparability of the analytical systems. NCCLS document EP9— *Method Comparison and Bias Estimation Using Patient Samples,* describes the procedures and factors laboratorians should consider in carrying out method comparison evaluations. The reader should refer to that document for details. In general, if the analytical system in question has similar imprecision and known interferences, uses the same or comparable standards or calibrators, reports in the same units, and is acceptably comparable in absolute values to the current method as judged by methods experimentation, then the reference interval can be transferred to the new or alternate analytical system. If, however, acceptable comparability is not validated as outlined in NCCLS document EP9, then a new reference value study must be undertaken.

If a clinical laboratory wishes to transfer a reference interval established by another laboratory or diagnostic test manufacturer for the same or acceptably comparable (as previously established) analytical system, 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.

8.2 Validation

Essentially three approaches can be considered to assess the acceptability of the transference of a reference interval for the same or acceptably comparable analytical system.

(1) The acceptability of the transfer may be rather subjectively assessed by inspection of the pertinent factors of the original appropriate reference value study. To be able to do this, all of the reference population demographics and geographics 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.

(2) 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 that they represent a statistically homogeneous group of results, i.e., that none of the results appears to be an outlier. To test for possible outliers, Reed's "one-third" rule cited earlier should be applied. Any apparent outliers should be discarded and new patient specimens obtained to replace them so that 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 2 of the 20 tested subjects' values (or 10% of the test results) fall outside those original reported limits. If three or more test results do 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, the user should re-examine 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 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 of false rejection of the donor 95% limits (i.e., the probability of rejecting the donor limits when at least 95% of the user's population do fall within these limits) is 5 to 7%, as close as one can come using discrete statistics to the conventional goal of 5% for the probability of false rejection of the null hypothesis.

(3) The acceptability of the transfer may also be assessed and validated by examining somewhat more reference individuals (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. Also if there are substantial differences 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.

The reference individuals are selected and the reference values are obtained in accordance with the previously discussed guidelines in Sections 4 and 5. 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 (Section 7). 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 subclassing protocol, additional sampling is required for further comparison, or a full-scale reference value study should be undertaken.

Others have suggested that there may be a more general solution to the transference problem, including between different analytical systems, whereby the “true” biological reference distribution is determined and then transferred.[20-22] One would adjust or correct the reference distribution of the donor laboratory for the effects of local methodological bias and imprecision to obtain an estimate of the "true" biological distribution. The converse adjustments are then made by the receiving laboratory for their respective analytical system to obtain an estimate of the reference distribution. This type of solution was thought to address transference between different methods, even if the units of measurement were different, provided an appropriate mathematical relationship between the two methods can be described. However, the approach has not been fully explored and requires additional research. It does not address, nor could it correct for, potential differences in the reference subject populations.

9 Presentation of Reference Values

9.1 Introduction

This section addresses the presentation of observed or patient values related to reference values.[5] The comments in this section are divided into two groups. Section 9.2 addresses the presentation of reference values by laboratories and end users. Section 9.3 covers the same subject as it applies to the manufacturers of quantitative clinical laboratory diagnostics tests.

9.2 Laboratory Presentation

Every quantitative clinical result should be accompanied by an appropriately presented reference interval. Lengthy reports that include the results of many tests should include some way of highlighting those results not within the reference interval. The reference intervals applied should reflect the subclass partitions that have been determined to be significant for that laboratory's particular reference population. It is helpful to flag patient results in a report indicating the relationship of the results relative to the reference interval. The term “reference interval” should be used rather than NORMAL, USUAL, or EXPECTED. Printing “HIGH” or “LOW” adjacent to a result is an acceptable option. Figure 4 is a representative example.

The use of forms with the reference interval preprinted requires that reference intervals for all appropriate subclasses be included and can result in a confusing report. 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 will be 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.

A document should be available to all users of a laboratory service, describing the reference population and the details of the reference intervals study. This document should be updated anytime a change is made in the laboratory that affects the reference intervals in use. A memo addressing changes in a reference interval should be sent to all users of the laboratory. Included should be details that indicate the number and demographics of the reference individuals, the assessment of health criteria used, the exclusion and partitioning criteria used with the reference sample, and the total size of the subclasses. The preanalytical details of the study should be documented, along with details of the analytical method, imprecision and inaccuracy, and the statistical method of analysis used. For subclass reference values with insufficient number of subjects (newborns, etc.), report the percentiles, number of observations, and sometimes all the values and entire range.

THE HOSPITAL
PATHOLOGY LABORATORY
ONE HOSPITAL
ANYWHERE, USA

Patient name: DOE, JANE
Patient ID#: 1234
Specimen ID#: 1001
Age: 25 years
Sex: Female
SID comment:
Location: SICU

Current date/time: 01/07/95/12:37
Date/time collected: 01/07/95/7:22
Drawn by: J.L.
Doctor: BROWN

Test Name	Results		Units	Reference Interval	Codes
Albumin	5.2	HIGH	g/dL	3.9—4.9	#
AspAT	6.0	LOW	U/L	16.3—39.7	#
Chloride	85	LOW	mmol/L	95—103	#
Glucose	75		mg/dL	70—106	#
Potassium	4.0		mmol/L	3.5—4.8	#
Sodium	150	HIGH	mmol/L	135—145	#
Phosphorus	4.0		mg/dL	2.4—4.7	#
Triglycerides	250	HIGH	mg/dL	35—160	#

Reviewed by:_____

Figure 4. Sample Laboratory Report

9.3 Manufacturer Presentation

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 9.2.

Manufacturers of instruments and reagents for quantitative diagnostic tests should provide reference interval information in the product labeling (operations manual and package inserts). In that labeling, manufacturers should be able to reference this document and have it understood that certain basic criteria of reference sample size, control of preanalytical and analytical variables, and statistical treatment have been met. 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. Manufacturers should indicate whether the most common partitioning factors have been examined for subclass differences, i.e., sex, age, fasting/nonfasting, time of day, pregnancy, and posture.

The reference population used by manufacturers should have a geographic distribution similar to the geographic distribution of their markets. It is important to recognize that there might be subclass differences in reference individuals from region-to-region that reflect not only geographical differences but also variables such as environment, diet, and ethnic background.

To support transference, all of the details of the reference interval study should be readily available on request. These details should be the same as those suggested in Section 9.2 for laboratories and include, in addition, the results for each of the reference individuals.

10 Other Issues

10.1 Qualitative Analysis

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

10.2 Therapeutic Drug Levels

The guideline does not address the determination of therapeutic drug levels. This requires a different study. The population required for these studies would necessarily have 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.

10.3 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 will be many other factors to consider in addition to all those that are of “routine” concern.

10.4 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 would be the referent. This would involve a separate study of the biological component for the total variance of observed values in each subject under given experimental conditions.[23]

10.5 "Critical Values"/Medical Decision Limits

This guideline is not intended to address the issue of "critical values" or other medical decision limits. Decision limits are different from reference limits, because they are based on other scientific and medical knowledge, and they may be related to a specific medical condition. Usually they are not derived in the same manner as reference intervals. In some instances, for example the National Cholesterol Education Program's risk-associated lipid limits, the medical decision limits may be more appropriate for use than population-based "healthy reference intervals." In these instances, the medical decision limits, rather than the "healthy reference intervals," should accompany the patient's results on the laboratory report.

10.6 Manufacturer's Data

Manufacturers should supply all reference value data and reference value study protocol information to clinical laboratories upon request. This may provide useful information, particularly for the more esoteric analytes, and minimize the additional work to be done to verify a reference interval. This data will also be helpful in assessing the transferability of reference intervals for the analyte(s) in question.

11 Summary

In this document, the subcommittee recommends a basic approach and a more systematic process for determining a reference interval. From the respective literature, a process is described that is reasonable and at the same time consistent with the principles for the production of reliable reference values. The basic principles that follow are uniformly important and must underlie any reference value study:

(1) The selection of the 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 process for evaluating health status is not rigorously defined—just as health cannot be rigorously defined. The process varies in elaborateness and expense depending on its intended use; however, the description allows other investigators or the end user to understand any limitations of the reference value study. The better the reference individuals are defined and described, the greater the value of the reference interval study. The subcommittee 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. 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 to be tested.

(3) The nonparametric method of estimation of the reference interval is recommended for its simplicity and reliability. More importantly, this method requires no specific assumption about the mathematical form of the probability distribution of reference values. More elaborate statistical techniques may be of value in particular instances, but these instances will require significant statistical expertise.

(4) A uniform process for detecting and discarding outlier values is recommended.

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

(6) The document recommends a minimum sample of 120 reference values for each reference population or subclass. This is the smallest number of samples that allows the determination of a 90% confidence interval around the reference limits (e.g., the 2.5th and 97.5th percentile). Greater confidence or improved precision in an estimated 95% reference interval can be accomplished from a larger sample of reference individuals.

(7) Transference of reference intervals may be accomplished only under certain conditions as outlined in the document. The acceptability of the transfer may be assessed by careful inspection of all of the details of the donor reference value for compatibility without any laboratory validation studies; or by comparison of a smaller reference sample validation study in the receiving laboratory to the donor laboratory's reference values using the shortened 20-sample approach; or by the standard normal deviate test (described for identifying the need for separate subclass reference intervals). However, if there are substantial differences in the demographics or geographics between the populations of the receiving and donor laboratories, transference is not recommended.

(8) The proper and adequate presentation of reference values and reference intervals to laboratories and to clinicians is defined. All details of a reference interval study should be readily available on request.

Finally, the NCCLS Subcommittee on Reference Intervals recognizes the need for a national or regional program for the collection of reference samples and advocates the establishment of specimen banks for reference interval studies. With appropriate financial support, well-defined protocols, and large sample populations, the quality of reference intervals would be significantly enhanced. Demographic and geographic differences in reference intervals would be detectable and available for use in clinical laboratory medicine. The cost of duplication of many local reference interval studies could be reduced, and the cost of the use of inappropriate reference intervals could be avoided. Reliable transference of reference intervals could be documented through the use of these banked specimens.

While a national bank of specimens would be optimal, it would also be useful to accumulate data on reference interval studies gathered under conditions consistent with this NCCLS guideline for a national data bank. We strongly encourage the consistent approach outlined and documentation of the details and variables examined to this end.

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

[8] 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.

[9] Hjelm M. Preparing reference individuals for blood collection. In: Grasbeck R, Alstrom T, eds. *Reference Values in Laboratory Medicine*. John Wiley and Sons Ltd.; 1981:105-108.

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

[11] Galen RS, Peters T. Analytical goals and clinical relevance of laboratory procedures. In: Tietz NW, ed. *Textbook of Clinical Chemistry.* Philadelphia: WB Saunders; 1986:387-409.

[12] Peters P, Westgard JO. Evaluation of methods. In: Tietz NW, ed. *Textbook of Clinical Chemistry.* Philadelphia: WB Saunders; 1986:410-423.

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

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

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

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

[17] 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.

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

[19] 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.

[20] PetitClerc C, Kelley A. Transferability of reference data. In: Grasback R, Alstrom T, eds. *Reference Values in Laboratory Medicine*. Chichester, England: John Wiley & Sons Ltd.; 1981.

[21] Strike PW, Michaeloudis A, Greer AJ. Standardizing clinical laboratory data for the development of transferable computer-based diagnostic programs. *Clin Chem.* 1986;32:222.

[22] Groth TL. Internal and external quality assurance. *Scand J Clin Lab Invest.* 1989;193(suppl 49):87-94.

[23] Winkel P. The use of the subject 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.

Summary of Comments and Subcommittee Responses

C28-A: *How to Define, Determine, and Utilize Reference Intervals in the Clinical Laboratory; Approved Guideline*

General

1. We believe it is important to emphasize the importance of regional validation and verification of subclass partitioning as essential to the utility of a given laboratory's use of published reference intervals. The concept of a national or geographically universal reference interval, while perhaps desirable, is an externally difficult achievement.

- **The committee agrees. Our statements suggested that there is value in national or regional programs for the collection of specimens for reference interval determinations by reducing the cost of such determinations and by "detecting demographic and geographic differences." We did not suggest that there is one universal reference interval.**

2. It would be helpful if manufacturers would be required to carry out these guidelines and then supply the information to their customers. The idea of a manufacturer data bank where laboratories could contribute data under specified criteria is an excellent one.

- **The committee agrees. It was our hope that diagnostic test manufacturers would adopt this protocol and reliably establish reference intervals for the customer laboratories. We are encouraged that this is happening by some of the industry.**

3. The procedure described would be cumbersome, labor-intensive, and costly to perform. If a shortened version could accomplish the same end, it would be more user-friendly in an actual work setting.

- **The current recommendation is as short as it can get and still produce reliable reference intervals. It was a major goal of the original subcommittee to make this as short and efficient a process as possible.**

4. It may be useful for NCCLS to have an accompanying document with several different examples. This way people not familiar with the statistics could better follow the logic to be applied, particularly with the three suggested approaches of transference.

- **The Area Committee on Clinical Chemistry and Toxicology agrees that this suggestion is a good idea and will consider future projects on transference.**

5. Why was calcium data expressed as mg/L, rather than the routine mg/dL?

- **All the units in the document have been revised to be expressed as mg/dL (mmol/L).**

6. We believe there are two needs not addressed by the current document: (1) laboratories which do not have enough specimens to perform the nonparametric analysis and (2) have samples which may contain significant numbers of specimens from an unhealthy population. The document does not address outlier detection. We suggest that the committee consider the methods of Harrell and Davis, as well as the use of a robust estimator to determine reference intervals. The robust estimator has been published by Horn, Copeland and Pesce in *Clinical Chemistry (A robust approach to reference interval estimation and evaluation; Clinical Chemistry* 44:3 622-631 (1998) and is an application of a well-established and accepted statistical method.

- **There appears to be merit to the approach of Harrell and Davis and the "Robust Approach" published by Horn, Copeland and Pesce, particularly with regard to the problem of having less than the recommended number of reference subjects. Because these approaches have not been tried, their validity at present is unknown. To accept them for establishing reliable reference intervals will require more discussion and a consensus. The Area Committee on Clinical Chemistry and Toxicology will consider this topic as a possible future NCCLS project at its next meeting and the Area Committee on Evaluation Protocols is currently developing a project on "quality goals for acceptable performance and threshold criteria for outliers."**

Section 2.2

7. The term "reference range" is not clearly defined while it is being differentiated from the term

- **This definition of reference range and that of the IFCC refers to the entire range of values, i.e., minimum and maximum values of the entire 100% set of values from the reference subjects. Whereas, the reference interval refers to the interval between two reference limits that includes, usually, 95% of the reference values.**

Section 5

8. The final paragraph seems to indicate that a laboratory might have several instruments performing the same test with these instruments giving different results, which would require multiple reference intervals for the same analyte within a single laboratory. A laboratory should adjust the instruments so that they give consistent values no matter where a sample is assayed. Slopes, intercepts, or other factors should be added to an instrument, and comparisons of the sample results or other statistical methods should be used to make sure all methods match.

- **The last sentence of that section has been modified to say that if the different methods or measuring systems cannot be made to be comparable, then separate reference intervals may have to be established. Unfortunately, there will be situations in our laboratory testing environment when different methods cannot be made comparable simply because they are not measuring the same substance(s) and/or do not have the same analytical specificity with regard to interferences.**

Summary of Delegate Comments and Responses

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

1. This document is much too complicated and expensive for its intended audience. This would be appropriate for manufacturers, but not end users.

- **The subcommittee originally responsible for development of C28-A worked carefully to find a balance in the text between the complex statistical approaches appropriate for manufacturer needs and a "user-friendly" approach for use by laboratorians and clinicians. The committee believed the resulting document described the minimum requirements for an adequate and appropriate reference interval determination. The area committee is reaffirming the approach contained in the revised C28-A2 document as evidenced by the response to Comment 3 in the Summary of Comments and Committee Responses: "The current recommendation is as short as it can get and still produce reliable reference intervals. It was a major goal of the original subcommittee to make this as short and efficient a process as possible."**

2. Although the document could stand as it is, the writing style is somewhat dense in the statistical sections (i.e., Section 7.3). Also, with the common availability of PCs and software tools, we could "raise the bar" and be more precise in our statistical testing without burdening the clinician with cumbersome calculations.

- **The project's management team believes that alternative statistical approaches may have merit. However, evaluation of the suggested approaches would require detailed evaluation and possible development of supporting software in order to be useful. At this time, the management team does not believe that such extensive changes in C28 are warranted; alternative approaches may be considered during the next revision of the document.**

Related NCCLS Publications*

C24-A2 **Statistical Quality Control for Quantitative Measurements: Principles and Definitions; Approved Guideline—Second Edition (1998).** This guideline provides definitions of analytical intervals, plans for quality control procedures; and guidance for quality control applications.

EP9-A **Method Comparison and Bias Estimation Using Patient Samples; Approved Guideline (1995).** This document addresses procedures for determining the bias between two clinical methods or devices, and for the design of a method comparison experiment using split patient samples and data analysis.

GP16-A **Routine Urinalysis and Collection, Transportation, and Preservation of Urine Specimens; Approved Guideline (1995).** This guideline describes routine urinalysis test procedures that address materials and equipment, macroscopic examinations, clinical analyses, and microscopic evaluations.

H3-A4 **Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture**; **Approved Standard—Fourth Edition (1998).** This document provides procedures for the collection of diagnostic specimens by venipuncture, including line draws, blood culture collection, and venipuncture in children. It also includes recommendations on order of draw.

H4-A4 **Procedures and Devices for the Collection of Diagnostic Blood Specimens by Skin Puncture; Approved Standard—Fourth Edition (1999).** This document provides a technique for the collection of diagnostic blood specimens by skin puncture, including recommendations for collection sites and specimen handling and identification. Specifications for disposable devices used to collect, process, and transfer diagnostic blood specimens obtained by skin puncture are also included.

H11-A3 **Procedures for the Collection of Arterial Blood Specimens; Approved Standard—Third Edition (1999).** This standard describes principles for collecting, handling, and transporting arterial blood specimens. This document is aimed at reducing collection hazards and ensuring the integrity of the arterial specimen.

H18-A2 **Procedures for the Handling and Processing of Blood Specimens; Approved Guideline (1999).** This document includes criteria for preparing an optimal serum or plasma sample and for the devices used to process blood specimens.

H21-A3 **Collection, Transport, and Processing of Blood Specimens for Coagulation Testing and General Performance of Coagulation Assays; Approved Guideline—Third Edition (1998).** This guideline contains procedures for collecting, transporting, and storing blood; processing blood specimens; storing plasma for coagulation testing; and provides general recommendations for performing the tests.

* Proposed- and tentative-level documents are being advanced through the NCCLS consensus process; therefore, readers should refer to the most recent editions.

Related NCCLS Publications (Continued)

M29-A **Protection of Laboratory Workers from Instrument Biohazards and Infectious Disease Transmitted by Blood, Body Fluids, and Tissue; Approved Guideline (1997).** A consolidation of M29-T2 and I17-P, this document provides guidance on the risk of transmission of hepatitis viruses and human immunodeficiency viruses in any laboratory setting; specific precautions for preventing the laboratory transmission of blood-borne infection from laboratory instruments and materials; and the recommendations for the management of blood-borne exposure.

NRSCL8-A **Terminology and Definitions for Use in NCCLS Documents; Approved Standard (1998).** This document provides standard definitions for use in NCCLS standards and guidelines, and for submitting candidate reference methods and materials to the National Reference System for the Clinical Laboratory (NRSCL).

NOTES

NOTES

NCCLS ▼ 940 West Valley Road ▼ Suite 1400 ▼ Wayne, PA 19087 ▼ USA ▼ PHONE 610.688.0100
FAX 610.688.0700 ▼ E-MAIL: exoffice@nccls.org ▼ WEBSITE: www.nccls.org ▼ ISBN 1-56238-406-6