EP14-A
Vol. 21 No. 3

Replaces EP14-P
Vol. 18 No. 2

Evaluation of Matrix Effects; Approved Guideline

This document provides guidance for evaluating the error or bias in analyte measurements that is due to the sample matrix (physiological or artificial) when two analytical methods are compared.

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.

Evaluation of Matrix Effects; Approved Guideline

Abstract

Evaluation of Matrix Effects; Approved Guideline (NCCLS document EP14-A) was developed for manufacturers, regulators, and providers of proficiency testing or external quality assessment programs, although it will also be useful to clinical laboratories as well. The document will help users to determine whether matrix effects are the source of unexpected results that are sometimes observed with processed samples when two analytical methods are compared; to identify and quantify the magnitude of the effects; and to assure that laboratory performance is evaluated fairly if matrix effects are present. The suggested protocols were developed using patient specimens as the standard of comparison. A list of definitions is included.

NCCLS. *Evaluation of Matrix Effects; Approved Guideline.* NCCLS document EP14-A (ISBN 1-56238-434-1). NCCLS, 940 West Valley Road, Suite 1400, Wayne, Pennsylvania 19087-1898 USA, 2001.

EP14-A
ISBN 1-56238-434-1
ISSN 0273-3099

Evaluation of Matrix Effects; Approved Guideline

Volume 21 Number 3

Fred D. Lasky, Ph.D., Chairholder
Donald Joe Boone, Ph.D.
John H. Eckfeldt, M.D., Ph.D.
Carolyn S. Feldkamp, Ph.D.
David J. Hassemer, M.S.
Sandy Krishnamurthy, M.S.
Herbert K. Naito, Ph.D., M.B.A.
Alan Posner, Ph.D.

Suggested Citation

(NCCLS. *Evaluation of Matrix Effects; Approved Guideline*. NCCLS document EP14-A [ISBN 1-56238-434-1]. NCCLS, 940 West Valley Road, Suite 1400, Wayne, Pennsylvania 19087-1898 USA, 2001.)

Proposed Guideline
April 1999

Approved Guideline
March 2001

ISBN 1-56238-434-1
ISSN 0273-3099

Committee Membership

Area Committee on Evaluation Protocols

Jan S. Krouwer, Ph.D. **Chairholder**	**Bayer Diagnostics** **Medfield, Massachusetts**
Carl C. Garber, Ph.D. **Vice-Chairholder**	**Quest Diagnostics, Inc.** **Teterboro, New Jersey**

Subcommittee on Matrix Effects

Fred D. Lasky, Ph.D. **Chairholder**	**Ortho-Clinical Diagnostics** **Rochester, New York**
Donald Joe Boone, Ph.D.	Centers for Disease Control and Prevention Atlanta, Georgia
John H. Eckfeldt, M.D., Ph.D.	Fairview-University Medical Center Minneapolis, Minnesota
Carolyn S. Feldkamp, Ph.D.	Henry Ford Hospital Detroit, Michigan
David J. Hassemer, M.S.	Wisconsin State Laboratory of Hygiene Madison, Wisconsin
Sandy Krishnamurthy, M.S.	Fullerton, California
Herbert K. Naito, Ph.D., M.B.A.	VA Medical Center Cleveland, Ohio
Alan Posner, Ph.D.	BioResource Technology Fort Lauderdale, Florida

Advisors

Randall L. Barbour, Ph.D.	SUNY Health Science Center at Brooklyn Brooklyn, New York
Prof. Dr. J. Buttner	Medizinische Hochschule Hannover Hannover, Germany
Dennis Collin, Ph.D.	Roche Diagnostic Systems Somerville, New Jersey
Bradley E. Copeland, M.D.	University of Cincinnati Medical Center Cincinnati, Ohio
John Costas	Medlab, Inc. New Castle, Delaware

Advisors (Continued)

Thomas E. England, Ph.D.	MDS Laboratories Ltd. Ontario, Canada
Kevin D. Fallon, Ph.D.	Instrumentation Laboratory Lexington, Massachusetts
Melvin R. Glick, Ph.D.	Wishard Memorial Hospital/Indiana University Medical Center Indianapolis, Indiana
Joan H. Howanitz, M.D.	New York, New York
Verlin K. Janzen, M.D.	Hutchinson Clinic, P.A. Hutchinson, Kansas
J. Mehsen Joseph, Ph.D.	Maryland Department of Health & Mental Hygiene Baltimore, Maryland
Harvey W. Kaufman, M.D.	Quest Diagnostics, Incorporated Teterboro, New Jersey
Arthur Kessner, Ph.D.	Beckman Coulter, Inc. Brea, California
Yun Sik Kwak, M.D., Ph.D.	VA Medical Center Cleveland, Ohio
Ronald H. Laessig, Ph.D.	University of Wisconsin Madison, Wisconsin
Glenda Lee	Tulane Medical Center New Orleans, Louisiana
Stanley S. Levinson	VA Medical Center, Louisville Louisville, Kentucky
Frank J. Liu, M.D.	M.D. Anderson Cancer Center Houston, Texas
Sharon Longhurst	Abbott Laboratories Abbott Park, Illinois
W. Gregory Miller, Ph.D.	Virginia Commonwealth University Richmond, Virginia
Nancy Moore	FDA Center for Devices and Radiological Health Rockville, Maryland
John D. Nagle, Ph.D.	CASCO•NERL Diagnostics Yarmouth, Maine

Advisors (Continued)

Robert M. Nakamura, M.D.	Scripps Clinic & Research Foundation La Jolla, California
David M. Obzansky, Ph.D.	Dade Behring, Inc. Elkton, Maryland
Michael J. Ondras	St. Margaret Memorial Hospital Pittsburgh, Pennsylvania
Donald R. Parker, Ph.D.	Bayer Corporation Elkhart, Indiana
Robert Rej, Ph.D.	Wadsworth Center for Laboratories and Research Albany, New York
John W. Ross, M.D.	Kennestone Hospital Marietta, Georgia
Brenda Warren, M.S., M.T.(ASCP)	St. Marys Hospital Medical Center Madison, Wisconsin
Alan Witheiler, Ph.D.	NHL Corp. La Jolla, California
Thomas R. King, M(ASCP), M.P.H. *Staff Liaison*	NCCLS Wayne, Pennsylvania
Patrice E. Polgar *Editor*	NCCLS Wayne, Pennsylvania
Donna M. Wilhelm *Assistant Editor*	NCCLS Wayne, Pennsylvania

Active Membership
(as of 1 January 2001)

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.
SmithKline Beecham Pharmaceuticals

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.
Asociacion de Laboratorios de Alta Complejidad
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
COLA
College of American Pathologists
College of Medical Laboratory Technologists of Ontario
College of Physicians and Surgeons of Saskatchewan
Fundación Bioquímica Argentina
International Association of Medical Laboratory Technologists
International Council for Standardization in Haematology
International Federation of Clinical Chemistry
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 Quality Management Program-Laboratory Service
RCPA Quality Assurance Programs PTY Limited
Sociedade Brasileira de Analises Clinicas
Sociedade Brasileira de Patologia Clinica
Sociedad Espanola de Quimica Clinica

Government Members

Association of Public Health Laboratories
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 Center for Veterinary Medicine
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
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.
Agilent Technologies, Inc.
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-Development SRL
Bio-Inova Life Sciences International
BioMedia Laboratories Sdn Bhd
bioMérieux, Inc.
Biometrology Consultants
Bio-Rad Laboratories, Inc.
Biotest AG
Bristol-Myers Squibb Company
Canadian External Quality Assessment Laboratory
Capital Management Consulting, Inc.
Checkpoint Development Inc.
Clinical Design Group Inc.
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
Eiken Chemical Company, Ltd.
Enterprise Analysis Corporation
Fort Dodge Animal Health
General Hospital Vienna (Austria)
Gen-Probe
Glaxo-Wellcome, Inc.
Greiner Meditech, Inc.
Health Systems Concepts, Inc.
Helena Laboratories
Home Diagnostics, Inc.
I-STAT Corporation
Instrumentation Laboratory
International Technidyne Corporation
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, Inc.
Merck & Company, Inc.
mvi Sciences (MA)
Nabi
Neometrics is.
Nichols Institute Diagnostics (Div. of Quest Diagnostics, Inc.)
Nissui Pharmaceutical Co., Ltd.
Nippon Becton Dickinson Co., Ltd.
Norfolk Associates, Inc.
Organon Teknika Corporation
Ortho-Clinical Diagnostics, Inc. (Raritan, NJ)
Ortho-Clinical Diagnostics, Inc. (Rochester, NY)
Oxoid Inc.
Pfizer Inc
Pharmacia & Upjohn
Premier Inc.
Procter & Gamble Pharmaceuticals, Inc.
The Product Development Group
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 Pharmaceuticals
SmithKline Beecham, S.A.
Streck Laboratories, Inc.
SurroMed, Inc.
Sysmex Corporation (Japan)
Sysmex Corporation (Long Grove, IL)
The Toledo Hospital (OH)
Trek Diagnostic Systems, Inc.
Vetoquinol S.A.
Visible Genetics, Inc.
Vysis, Inc.
Wallac Oy
Wyeth-Ayerst
Xyletech Systems, Inc.
YD Consultant
Yeongdong Pharmaceutical Corporation

Trade Associations

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

Associate Active Members

20th Medical Group (SC)
67th CSH Wuerzburg, GE (NY)
121st General Hospital (CA)
Academisch Ziekenhuis-VUB (Belgium)
Acadiana Medical Laboratories, LTD (LA)
Adena Regional Medical Center (OH)
Advocate Laboratories (IL)
The Aga Khan Hospital & Medical College, Karachi (Pakistan)
Albany Medical Center Hospital (NY)
Albemarle Hospital (NC)
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)
Arkansas Department of Health
Armed Forces Research Institute of Medical Science (APO, AP)
Associated Regional & University Pathologists (UT)
Aurora Consolidated Laboratories (WI)
Bay Medical Center (MI)
Baystate Medical Center (MA)

Bonnyville Health Center (Alberta, Canada)
Boulder Community Hospital (CO)
Brantford General Hospital (Ontario, Canada)
Brazileiro De Promocao (Brazil)
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)
Chang Gung Memorial Hospital (Taiwan)
Children's Hospital (LA)
Children's Hospital (NE)
Children's Hospital & Clinics (MN)
Children's Hospital King's Daughters (VA)
Children's Hospital Medical Center (Akron, OH)
Children's Hospital of Philadelphia (PA)
Clarian Health–Methodist Hospital (IN)
Clendo Lab (Puerto Rico)
CLSI Laboratories (PA)
Columbus County Hospital (NC)
Commonwealth of Kentucky
CompuNet Clinical Laboratories (OH)
Covance Central Laboratory Services (IN)
Danish Veterinary Laboratory (Copenhagen, Denmark)
Danville Regional Medical Center (VA)
Deaconess Hospital (MO)
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)
Dynacare Memorial Hermann Laboratory Services (TX)
E.A. Conway Medical Center (LA)
Eastern Maine Medical Center
East Side Clinical Laboratory (RI)
Elyria Memorial Hospital (OH)
Emory University Hospital (GA)
Esoterix Center for Infectious Disease (TX)
Fairfax Hospital (VA)
Fairview-University Medical Center (MN)
Foothills Hospital (Calgary, AB, Canada)
Fort St. John General Hospital (Fort St. John, BC, 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)
Geisinger Medical Center (PA)
Grady Memorial Hospital (GA)
Guthrie Clinic Laboratories (PA)
Harris Methodist Fort Worth (TX)
Harris Methodist Northwest (TX)
Hartford Hospital (CT)
Headwaters Health Authority (Alberta, Canada)
Health Alliance Laboratory (OH)
Health Network Lab (PA)
Health Sciences Centre (Winnipeg, MB, Canada)
Heartland Health System (MO)
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 State Board of Health
Indiana University
Instituto Scientifico HS. Raffaele (Italy)
International Health Management Associates, Inc. (IL)
Jersey Shore Medical Center (NJ)
Joel T. Boone Branch Medical Clinic (VA)
John C. Lincoln Hospital (AZ)
John Peter Smith Hospital (TX)
John Randolph Hospital (VA)
Kaiser Permanente (CA)
Kaiser Permanente (MD)
Kantousspital (Switzerland)
Kenora-Rainy River Regional Laboratory Program (Ontario, Canada)
Kern Medical Center (CA)
King Fahad National Guard Hospital (Saudi Arabia)
Kings County Hospital Center (NY)
Klinicni Center (Slovenia)
LabCorp (NC)
Laboratoire de Santé Publique du Quebec (Canada)
Laboratories at Bonfils (CO)
Laboratório Fleury S/C Ltda. (Brazil)
Laboratory Corporation of America (MO)
LAC and USC Healthcare Network (CA)
Lakeland Regional Medical Center (FL)
Lancaster General Hospital (PA)
Langley Air Force Base (VA)
LeBonheur Children's Medical Center (TN)
Lewis-Gale Medical Center (VA)
Libero Instituto Univ. Campus BioMedico (Italy)
Licking Memorial Hospital (OH)
Louisiana State University Medical Center
Magee Womens Hospital (PA)
Magnolia Regional Health Center (MS)
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 College of Virginia Hospital

Medicare/Medicaid Certification, State of North Carolina
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)
Mercy Medical Center Des Moines (IA)
Mescalero Indian Hospital (NM)
Methodist Hospitals of Memphis (TN)
Michigan Department of Community Health
Mississippi Baptist Medical Center
Monte Tabor – Centro Italo - Brazileiro de Promocao (Brazil)
Montreal Children's Hospital (Canada)
Montreal General Hospital (Canada)
Morton Plant Mease Health Care (FL)
Mount Sinai Hospital (NY)
Mount Sinai Medical Center (FL)
MRL Reference Laboratory (CA)
National University Hospital (Singapore)
Naval Surface Warfare Center (IN)
New Britain General Hospital (CT)
New England Fertility Institute (CT)
New England Medical Center Hospital (MA)
New York Hospital Medical Center of Queens
New York State Department of Health
NorDx (ME)
North Carolina Laboratory of Public Health
North Mississippi Medical Center
North Shore – Long Island Jewish Health System Laboratories (NY)
Northridge Hospital Medical Center (CA)
Northwestern Memorial Hospital (IL)
Ohio Valley Medical Center (WV)
Olin E. Teague Medical Center (TX)
O.L. Vrouwziekenhuis (Belgium)
Ordre professionnel des technologists médicaux du Québec
Ospedali Riuniti (Italy)
The Ottawa 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)
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)
Prodia Clinical Laboratory (Indonesia)
Providence Health System (OR)
Providence Seattle Medical Center (WA)
Queen Elizabeth Hospital (Prince Edward Island, Canada)
Queensland Health Pathology Services (Australia)
Quest Diagnostics, Incorporated (AZ)
Quest Diagnostics Incorporated (CA)
Quintiles Laboratories, Ltd. (GA)
Regions Hospital
Research Medical Center (MO)
Rex Healthcare (NC)
Rhode Island Department of Health Laboratories
Riyadh Armed Forces Hospital (Saudi Arabia)
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 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)
St. Mary's Hospital & Medical Center (CO)
Ste. Justine Hospital (Montreal, PQ, Canada)
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)
South Bend Medical Foundation (IN)
Southern California Permanente Medical Group
South Western Area Pathology Service (Australia)
Speciality Laboratories, Inc. (CA)
Stanford Hospital and Clinics (CA)
State of Washington Department of Health
Stormont-Vail Regional Medical Center (KS)
Sun Health-Boswell Hospital (AZ)
Sunrise Hospital and Medical Center (NV)
Tenet Odessa Regional Hospital (TX)
Touro Infirmary (LA)
Tri-City Medical Center (CA)
Trident Regional Medical Center (SC)
Tripler Army Medical Center (HI)
Truman Medical Center (MO)
UCSF Medical Center (CA)
UNC Hospitals (NC)
Unilab Clinical Laboratories (CA)
University Hospital (Gent) (Belgium)

University Hospital (TX)
The University Hospitals (OK)
University of Alberta Hospitals (Canada)
University of Chicago Hospitals (IL)
University of Florida
University of Medicine & Dentistry, NJ University Hospital
University of the Ryukyus (Japan)
University of Texas M.D. Anderson Cancer Center
University of Virginia Medical Center
University of Washington
UPMC Bedford Memorial (PA)
UZ-KUL Medical Center (Belgium)
VA (Dayton) Medical Center (OH)
VA (Denver) Medical Center (CO)
VA (Kansas City) Medical Center (MO)
VA (Martinez) Medical Center (CA)
VA (San Diego) Medical Center (CA)
VA (Tuskegee) Medical Center (AL)
VA Outpatient Clinic (OH)
Vejle Hospital (Denmark)
Virginia Department of Health
Viridae Clinical Sciences, Inc. (Vancouver, BC, Canada)
Warde Medical Laboratory (MI)
Washoe Medical Center Laboratory (NV)
Watson Clinic (FL)
Wilford Hall Medical Center (TX)
William Beaumont Hospital (MI)
Williamsburg Community Hospital (VA)
Winn Army Community Hospital (GA)
Wishard Memorial Hospital (IN)
Yan Chai Hospital (P.R. China)
Yonsei University College of Medicine (Korea)
York Hospital (PA)
Zale Lipshy University Hospital (TX)

Contents

Foreword

The presence of matrix effects in analytical systems used in the clinical laboratory has been a source of serious concern for many years. Although, in the literature, there are many references to the apparent incompatibility of fluids and analytical methods, when this work was first proposed there were no generally accepted guidelines that demonstrate how to identify and quantify the magnitude of the biases. Because these effects are commonly observed in external quality assessment schemes (EQAS) or proficiency testing (PT) survey results, protocols are needed to determine the presence or absence of these effects. Only then can the laboratorian assess whether the observed bias will have an impact on patient care.

Determining the presence or absence of matrix effects allows users, manufacturers, and those responsible for evaluating EQAS and PT data to distinguish between a true malfunction of the system and incompatibility between the system and the material being tested. The real difference is that analytical system malfunctions affect patient care, while matrix effects limit how the analytical system can be evaluated and monitored. When matrix effects are present with system calibrators, calibrator values should be adjusted so that reported patient results are not affected. In fact, this has become standard practice among manufacturers, and obviates the need for searching for reference materials with commutability; therefore, commutability will not be discussed further. For further discussion, see reference 1.

The Subcommittee on Matrix Effects was faced with a practical dilemma of definition. If a difference in results between methods is observed with processed samples using these protocols, an interfering substance is present. However, its source is not known in this early evaluation stage; it could be caused by a specific substance(s) or by the matrix— the milieu of the sample that differs from the specimens for which the method was designed. We decided for the purposes of this document to use the broadest interpretation; that is, this procedure is an effective way to identify whether an unexpected difference in results is observed in processed samples, and we direct the user to EP7 to test the source of the bias and quantify its magnitude in terms of analyte and interfering (if known) concentrations.

The subcommittee believes these protocols and the supporting information will be most useful to manufacturers and providers of external evaluation programs. Our objective is to provide ways to identify the presence of matrix effects so that improvements in method specificity and fluid compatibility (controls and calibrators) can be made, and to provide government regulators with a mechanism that can be used to distinguish between laboratories that are doing acceptable work from those that need improvement (based on the results of PT surveys). The subcommittee anticipates that this guideline will be helpful when differences in results between methods are observed with control or proficiency test materials that might affect an understanding of method performance.

The general rationale used to develop each protocol was that clinical laboratory methods are designed and developed to work optimally with patient specimens. Characteristics of manufactured control or calibrator materials that deviate significantly from the way patient specimens behave in analytical systems, with whatever response characteristics are used for measurement, can be called “matrix effects” because the source of the difference has not been identified. Pragmatically, for this document, an observed difference of unknown source is called a “matrix effect,” while a difference due to an identifiable substance or physical characteristic is an “interfering” (see Appendix A) and the user is referred to EP7—*Interference Testing in Clinical Chemistry*. Definitions are streamlined to account for known and unknown interferences.

Foreword (Continued)

The limitations of these procedures include (but are not limited to) the following:

- Subtle analytical differences that occur with consistency between different methods for measuring a given analyte may not be easily detectable. These protocols may not be sufficiently powerful to detect or identify the presence of these analytical differences. (Procedures described in Section 6.3(6) or 6.4(2) could be helpful.)

- No attempt is made to determine the clinical or regulatory significance of the magnitude of difference or bias between methods. The magnitude of the bias or difference might be used to compare to independently derived clinical or regulatory (e.g., PT) limits.

- These protocols cannot determine which of the two methods is more specific for measuring an analyte in a particular fluid.

- These protocols might not be usable within all disciplines of clinical analysis. The subcommittee continues to seek techniques that work within other areas.

Lastly, elimination of matrix effects requires either an improvement in the analytical specificity of methods or in the materials used for quality control, calibration, and/or external assessment. The clinical laboratory testing community should not lose sight of the fact that, in a "perfect" world, there would be no "matrix effect." In such a world, every routine method would have sufficient analytical specificity to produce accurate results with any fluid or material. This lack of analytical specificity, however, is the reason this guideline is needed.

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 for 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

Analytical interference, bias, matrix, matrix effect, physicochemical interference

Evaluation of Matrix Effects; Approved Guideline

1 Introduction

The interest in accuracy of testing in biological fluids has grown among the medical and laboratory professional community, as well as with the public. Regulations are in place that are meant to enhance the accuracy of the testing process. There is renewed emphasis on the use of external quality assessment schemes (EQAS) and proficiency testing (PT) to evaluate and monitor the accuracy of testing in clinical, reference, and physician's office laboratories.

Current scientific data suggest that such use of EQAS results is not always feasible because of matrix effects, which exist with many external control materials. These processed materials (including quality control and calibrating materials) sometimes do not behave like the fresh specimens routinely analyzed in the laboratory. Biases not generally seen with fresh biological fluids are frequently seen with EQAS, control, and calibrator materials. Because of these matrix effects, evaluating laboratory performance for accuracy of testing using PT can lead to inaccurate conclusions and, potentially, inappropriate regulatory sanctions.

Matrix effect phenomena involve the interplay of four major components in analytical testing: instrument design; reagent formulation; method principle; and control, calibrator, and PT material composition and processing technique. Within each of these categories are factors that contribute to the magnitude and direction (positive or negative) of the bias. The interactions that cause these matrix effects are complex and differ by discipline (e.g., chemistry, hematology) and by the nature of the materials used to calibrate and monitor performance of each method. The performance characteristics of a cellular suspension would be expected to differ from those of a protein-free filtrate.

Research is needed to characterize these interfering factors so that instruments, reagents, and fluids can be designed to minimize them. Until then, standardization of methods, as well as assessing or monitoring the accuracy of laboratory testing, will be difficult.

This document is complementary to EP7— *Interference Testing in Clinical Chemistry.* They are similar in that both provide protocols to help identify sources of error that can affect patient care and/or assess the suitability of a method. They are distinguishable in the following areas:

- EP14 focuses primarily on the difference between processed samples and patient specimens, while EP7 concentrates on how specific substances or conditions, e.g., the presence of an interfering substance or a change in viscosity, alter results in patient specimens.

- To evaluate the effect of interferences, EP14 compares performance of processed samples to a population of patient specimens, whereas EP7 uses criteria based on the precision of the method and the intra-individual variability of the analyte in the presence of increasing amounts of the interfering.

- The criteria used to determine if an effect is present are based on the dispersion of results from the patient specimens about the line of best fit in EP14, whereas EP7 uses the uncertainty of replicate measurements of a series of related pools that contain differing, known amounts of the substance (or change in condition) being investigated.

- EP14 compares the relative bias of processed samples to that of patient specimens, while the objective of EP7 is to quantify the observed bias as a function of the concentration of the interfering substance (or other characteristic) at specified concentrations of analyte.

2 Scope

This guideline is intended for diagnostic test manufacturers, external quality control and proficiency testing providers, and regulatory agencies. Although clinical laboratory use will probably be limited, because of the complexity of the calculations, the observations and conclusions should be useful to all professionals. The guideline provides protocols that evaluate matrix effects in processed samples that are used as standards, calibrators, controls, and EQAS or PT materials.

EP14 will assist in the education of clinical laboratorians, regulators, diagnostic manufacturers, and the public about the impact of matrix effects on the assessment of the quality of laboratory performance. For example, readers are warned that matrix effects, caused by the interaction of processed material and the analytical system, may suggest that erroneous results are being generated when in fact, the results are acceptable. Conversely, "acceptable" control results may also give a false sense of confidence that analytical systems are performing adequately. Terms and concepts used to report these and related issues are defined within this document.

The guideline can be used by laboratorians performing quantitative tests for a wide variety of analytes across various disciplines. The testing protocols attempt to accommodate situations where reference methods do not exist.

The protocols help laboratorians distinguish between effects caused by system malfunctions and those caused by use of processed samples. However, the protocols do not describe approaches that specifically establish the exact mechanism of the matrix effect(s).

By following the protocols, manufacturers and EQAS and PT providers should be able to provide some documentation to government or accrediting agencies on matrix effects to help avoid false conclusions about the adequacy of patient testing.

3 Precautions

Standard precautions should be followed when collecting or handling body fluids of any type. Please refer to the most current edition of NCCLS document M29— *Protection of Laboratory Workers from Instrument Biohazards and Infectious Disease Transmitted by Blood, Body Fluids, and Tissue.* Specimens from any patient may be infected with viruses such as human immunodeficiency virus (HIV) and hepatitis B virus (HBV). Proper techniques for collection of blood and other body fluids should be used to minimize risk to other patients, the laboratory staff, and all other hospital or clinical personnel. Gloves must be worn when handling test specimens.

The identification of matrix effects requires the use of controls, calibrators (standards), and materials used in external quality assessment schemes, as well as patient specimens, because it is necessary to compare results obtained with these materials. Most control and calibration materials have been treated to denature HIV and HBV, but they should still be handled with the same precautions as patient specimens. Extensive pipetting of these materials may be necessary for making dilutions. These samples should never be pipetted by mouth. Pipetting aids are available for every task. Bulbs or other suitable suction devices must always be used with pipets.

Controls, calibration materials, or diluents may contain azide, which is toxic. Azide may also form explosive compounds if it comes in contact with copper and lead plumbing. Products that contain azide should be flushed with excess water upon disposal down drains.

Materials for microbiological analysis should be handled in strict accord with the accepted techniques used to prevent the spread of the suspected organisms. Isolation hoods and sterile techniques should be used when indicated. Care should be taken to avoid forming aerosols. Because controls and standards of bacterial and viral assays may contain viable organisms, these should be handled appropriately.

4 Definitions[a]

Definitions are provided as they apply to this document. Some differ from other NCCLS documents because of the pragmatic requirements of these protocols. The use of a hierarchical approach to the source of observed biases (interferences) is illustrated in Appendix A, which should be used with definitions listed below.

Analyte, *n* **–** A substance or constituent for which the laboratory conducts testing; **NOTE:** This includes any element, ion, compound, substance, factor, infectious agent, cell, organelle, activity (enzymatic, hormonal, or immunological), or property the presence or absence, concentration, activity, intensity, or other characteristics of which are to be determined.

Analytical interference, *n* - **1)** The effect of a substance, either identified or unidentified, that causes a difference in the measured concentration or activity from the true value[2]; **NOTE:** The difference is systematically related to the concentration of the analytical interfering[3]; **Mechanistic interference**, *n* - **2)** A condition (of a substance or environment) that causes a difference between the population mean of the test results and an accepted reference value due to a change in the reaction mechanism[2]; **NOTE:** An example is an inhibitor that affects reagent system enzymes; **Physicochemical interference**, *n* - **3)** An environmental or structural difference from that of the patient specimens that causes a difference between the population mean of the test results and an accepted reference value due to a change in the measured physical chemical property[2]; **NOTE:** This is what has commonly been referred to as "matrix effect"; examples include the effect of different protein matrices on bilirubin spectra or the impact of proteins and lipids on the measurement of electrolytes in plasma in direct ion-selective electrode systems.

Average bias, *n* - The difference between the mean of the results of replicate measurements of a sample obtained using an analytical method, and those of an accepted reference or comparative method[1]; **NOTE:** In determining bias, the effect of imprecision is reduced by taking the average of a set of replicate test results.[2]

Calibration, *n* - 1**)** The process of testing and adjustment of an instrument, kit, or test system, to provide a known relationship between the measurement response and the value of the substance being measured by the test procedure. **2)** the set of operations that establish, under traceable conditions, the relationship between values indicated by a measuring instrument or measuring system for an established reference material and the corresponding value of a candidate reference material ***(WHO-BS/95.1793)***; **3)** The set of operations that establish, under specified conditions, the relationship between values of quantities indicated by a measuring instrument or measuring system, or values represented by a material measure or a reference material, and the corresponding values realized by standards. **NOTES:** a) the result of a calibration permits either the assignment of values of measurands to the indications or the determination of corrections with respect to indications; b) a calibration may also determine other metrological properties, such as the effect of influence quantities.

[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.

Commutability, *n* - The ability of a material to show interassay properties comparable with those of patient specimens; **NOTE:** This is a desired property for Certified Reference Material that is to be used as a standard for more than one method (refer to NCCLS document NRSCL13—*The Reference System for the Clinical Laboratory: Criteria for Development and Credentialing of Methods and Materials for Harmonization of Results*).

Comparative method, *n* - The method used as the basis for comparing two methods, e.g., in the evaluation of matrix effects; **NOTE:** The more specific this method is, the better the conclusion with regard to the source of the observed interference.

Evaluated method, *n* - That method for general clinical use that is being evaluated for a possible matrix effect.

Isoform, *n* - One of several forms of a single protein that have the same antigenic structure but that differ in minor amino acid content and/or steric structure.

Matrix, *n* - All components of a material system, except the analyte.

Matrix effect, *n* – **1)** The influence of a sample property, other than the analyte, on the measurement, and thereby on the value of the measurand; **2)** The physicochemical effect(s) (e.g., interference) of the matrix on the analytical method's ability to accurately measure an analyte.

Observed response, *n* - The measured physical or chemical parameter used to identify or quantify an analyte in comparison to an appropriate calibration system; **NOTE:** The observed response may be used by a system's internal processor and, therefore, the value is often not available to the testing personnel; examples include absorbance units, radioactive counts, and millivolt readings.

Processed sample , *n* - For the purposes of this document, a sample that is prepared to be used to mimic patient specimens; **NOTE:** a) It is considered a processed sample if it has been modified in any way that causes it to be different from fresh patient specimens, for example, freezing, lyophilization, adding nonendogenous substances, stabilizers, etc.; b) For this evaluation, these are the materials being evaluated for matrix effects.

Proficiency testing //external quality assessment, PT//EQA, *n* - A program in which multiple specimens are periodically sent to members of a group of laboratories for analysis and/or identification; in which each laboratory's results are compared with those of other laboratories in the group and/or with an assigned value, and reported to the participating laboratory and others.

Reported result, *n* - The value, in appropriate units, that is reported to the healthcare provider.

Residual, *n* - The difference between a given data point and its predicted value.

5 Principle of Evaluation

The evaluation of a matrix effect is based on the principle that the relationship between an observed response and the actual activity or concentration is dependent on the environmental conditions (e.g., temperature or matrix) at the time of measurement.[4] Because no measuring technique is completely specific, the observed relationship between any two methods will depend on the choice of the samples selected for comparison.[4] For clinical laboratory analysis, methods are designed to measure the concentration or activity in patient specimens, and a representative set of these specimens is used as the standard of comparison.

The magnitude of the matrix effect is evaluated by comparison to the "scatter" of results from the two methods being compared using a representative sample of patient specimens. The more heterogeneous the specimens, in terms of the interfering substance present, the larger the scatter expected in the data.

The magnitude of the bias measured in the processed sample(s) is compared to the resultant scatter of the patient specimens. This residual scatter represents the uncertainty of measurement of the evaluated method due to two factors: imprecision and nonspecificity. (The regression techniques used in these protocols use the assumption that there is no error in the comparative method represented on the x-axis.) The contribution of imprecision is reduced by replicate measurements; therefore, in these analyses, the primary contributor to scatter is the inherent interferences due to substances that are known or unknown (here called a "matrix effect"). The range of the scatter is represented by the prediction interval, which estimates the nonspecificity of the evaluated method for all patient specimens that would be tested. It is then possible to assert with reasonable probability whether the processed sample can be used to represent the set of patient specimens for the analyte being measured[5]; if the processed sample(s) result(s) is outside the prediction interval, a matrix effect is present.

Additionally, if a series of processed samples are related (as is often the case in an EQAS or PT event), such as being prepared from admixtures of common pools, regressing the results of these samples and comparing the line of best fit to the prediction interval can be used as a means of evaluation. This technique is especially helpful if the matrix effect bias is within the prediction interval, but is consistent or shows a relationship across all related processed samples.

Any conclusions from the study are limited to the specific variables of the processed samples, e.g., manufactured batch, sources of analytes used to supplement the sample, types of stabilizers that might be present, etc. Follow-up studies might be required to determine the source(s) of the observed biases.

6 Protocols

6.1 Materials

The following materials are needed for these protocols:

- evaluated method reagents, calibrators, and instrument system.

- comparative method reagents, calibrators, and instrument system. Use a method expected to show little or no matrix effect with processed calibrator or control samples. In order of preference, the comparative methods should fit the following descriptions, for example:

  — a National Reference System (NRSCL) definitive method (e.g., isotope dilution-mass spectrometry method for cholesterol);

  — a NRSCL reference method (e.g., the Abell-Kendall method for cholesterol);

  — an NRSCL-approved designated comparative method;

  — a commonly used method for the particular analyte in question.

  NOTE: Although ideally the comparative method should be free of matrix effects, this cannot be an absolute requirement. For practical reasons, a frequently used clinical method may be selected as the comparative method. When matrix effects are detected, however, the information obtained from these protocols will merely indicate that patient specimens and processed control fluids do not yield comparable results when used to measure a particular analyte with both methods. These

protocols will not identify whether the evaluated method or the comparative method has better specificity.

- processed samples, e.g., calibrators, control samples being studied.

- twenty fresh patient specimens with analyte concentrations or activities that are approximately evenly distributed over the concentration range of the processed samples of interest. Select patient specimens that are typically used for analysis (e.g., from both healthy and ill patients), and avoid those that are considered inappropriate for analysis due to the presence of known interferences. Frozen specimens may be included if freezing does not affect the measurements of either method.

6.2 Procedure

(1) Prepare the processed sample as directed.

(2) Using the evaluated method, analyze as a single analytical batch the 20 fresh patient specimens, with processed samples randomly interspersed between the fresh patient specimens. Repeat this process twice (sequential batches on a single day are preferable to eliminate the potential of shifts or drifts that can confound the data), preferably with separate calibrations. This yields three analytical results for each of the 20 patient specimens and the processed samples (see Appendix B). Perform a check for outliers, as recommended in NCCLS document EP9— *Method Comparison and Bias Estimation Using Patient Samples.*

(3) Using the comparative method, analyze (as a single analytical run or batch) the 20 fresh patient specimens, with the same processed samples randomly interspersed between patient specimens. Analyze the fresh specimens and processed samples at the same time as the evaluated method analyses. Repeat this process twice, preferably with separate calibrations. Perform an outlier check as recommended in EP9. If simultaneous analysis is not possible, information should be available to demonstrate that the comparative method results are not changed by the storage conditions used for the fresh patient specimens and for the processed samples.

(4) Freeze (preferably at -70 ºC) the 20 patient specimens and processed samples for future analysis. If any questions arise during or after data analysis, the specimens may be reanalyzed using another comparative method (e.g., an NRSCL definitive or reference method). Keep in mind that freezing may introduce a matrix effect by altering binding proteins, enzyme conformation, etc.

6.3 Data Analysis

As often occurs in statistical analysis, the user is asked to judge the utility and appropriateness of the statistical test for each data set. In these analyses, linearity, heteroscedacity, and each method's imprecision could affect interpretation of results. Use of incorrect assumptions will result in more difficulty in identifying the presence of a matrix effect; the prediction interval from the patient specimen set will be wider. We remind the user to keep in mind the intended purposes of each study. Standard statistical textbooks can be referenced.

(1) Plot the means of replicates of the 20 fresh patient specimens and the processed sample(s) (using different symbols) with the evaluated method results on the *y* axis and the comparative method on the *x* axis.

(2) Examine the distribution of the means of results from the fresh patient specimens obtained using the evaluated and comparative methods and verify the following prerequisites:

- Linear Regression Analysis

— The appearance of a linear relationship between the evaluated method and the comparative method results from patient's specimens without any noticeable curvature.

— The scatter in the *y*-direction around the regression line appears constant across the concentration range examined.

— Check the appropriateness of the data for regression analysis (refer to the most current edition of NCCLS document EP9—*Method Comparison and Bias Estimation Using Patient Samples*).

Then perform linear regression analysis using the means of the evaluated method results (from patient specimens) as the *y* value and the means of the comparative method results (from patient specimens) as the *x* value (see Appendix C, Example 1).

- Polynomial Regression Analysis:

If there appears to be curvature in the scatterplot of results obtained when comparing the evaluated method results and the comparative method results for the fresh patient specimens, perform a second-order polynomial regression analysis using the means of the evaluated method results as the *y* value and the means of the comparative method results as the *x* value (see Appendix C, Example 2)

$$y = a_o + a_1 x + a_2 x^2 \quad (1)$$

If the a_2 term in the second-order polynomial regression model is statistically different from zero (i.e., $p < 0.05$ determined by t-test), use the second-order regression model. If a_2 is not statistically different from zero (i.e., $p > 0.05$), use the linear regression model.

NOTE: Caution is advised to declare a method nonlinear based on 20 patient specimens, depending upon the distribution of their concentrations. Adding specimens to the experiment might be advisable. If linear regression analysis is performed (rather than polynomial regression) for methods that are not linearly related, it will be more difficult to determine whether or not a matrix effect is present. This is because the prediction interval for patient specimens will be wider if linear regression is used for nonlinear relationships. However, since linear regression analysis is usually more convenient, and if the matrix effects are large, starting with linear regression might satisfy the needs of the user.

NOTE: If the linear regression model is used when the relationship between the two methods is in fact curvilinear, the calculated prediction interval about the line will be broader. Therefore, it will be more difficult to identify a real difference (a matrix effect) between processed samples and patient specimens.

(3) If the scatter of patient specimen results around the regression line seems to increase in proportion to the analyte concentration, rather than being constant across the concentration range (i.e., variance divided by concentration is constant, rather than variance itself being constant across analyte concentration), perform a $\log_{10}$ transformation of the means of the evaluated method and comparative method results or, alternatively, plot results on log/log graph paper. Proceed through steps 1, 2, and 3 above; however, plot and perform linear or second-order polynomial regression analysis on the $\log_{10}$-transformed means, instead of the means themselves. To effectively estimate

the magnitude of scatter around the regression line, the variance of that scatter, or a transformation of the variance, must be constant.

(4) Using the formula shown below, compute the two-tailed 95% prediction interval for the mean of the fresh patient specimen *y* value at a given *x* value about the least-square linear regression line, the second-order polynomial regression line, or the $\log_{10}$-transformed (for *y* variable) regression line.

$$\overline{Y}_{pred} \pm t(0.975, n-2)\, S_{y \bullet x} \left[1 + \frac{1}{n} + \frac{(\overline{X}_i - \overline{\overline{X}})^2}{\sum(\overline{X}_i - \overline{\overline{X}})^2}\right]^{1/2} \quad (2)$$

where:

$\overline{Y}_{pred}$ = the predicted value of *y* at X_i based on an estimated regression curve;

n = the number of fresh patient specimens (not total number of replicates);

$S_{y \bullet x}$ = the standard error of regression =

$$\left[\sum (Y_{pred} - \overline{Y}_i)^2 / (n-2)\right]^{1/2};$$

$\overline{X}_i$ = *i*th value on the *x* axis (comparative method mean);

$\overline{Y}_i$ = *i*th value on the *y* axis (evaluated method mean); and

$\overline{\overline{X}}$ = the overall grand mean of the reference method means.

Compare each individual processed sample's mean *y* result to the statistically defined limits (95% prediction interval) derived from all patient specimen data using the equation. The user is reminded that if large differences exist in specificity of the methods used, poor correlation (large prediction interval) will result, making this procedure less or not effective. See examples in the appendix.

(5) Compare the magnitude of the processed sample deviation from the regression line (fresh patient specimens) to the 95% prediction intervals on the graph as illustrated in Appendix C. A matrix effect is present if the result of the processed sample(s) lie outside the prediction interval (Appendix C, Example 1). If the processed sample result is within the prediction interval, a matrix effect is probably absent (Appendix C, Example 2). However, if a persistent bias is observed among a set or group of processed samples and some or all of the biases are within the prediction interval, a matrix effect cannot be ruled out. If these sample sets are known to be related, such as admixtures of the same master pools, use the procedure described in Section 6.4(2) to aid in evaluation. The "allowable" or "acceptable" limit of the residual at any concentration would be evaluated against independent criteria.

NOTE: Caution is advised because PT limits published in the CLIA regulations are for single measurements of controls, while these protocols recommend using the mean value of triplicate assays. Therefore, a single measurement, performed for PT purposes, is not of equivalent experimental design, and therefore is not recommended to evaluate whether or not a matrix effect is present.

NOTE: Matrix effects that are statistically significant might not be clinically or quantitatively important if a control sample is being evaluated. However, a matrix effect of similar magnitude might be of concern if the processed sample were to be used as a calibrator.[6]

(6) **Groups of Interrelated Processed Samples.** If a group of processed samples are interrelated and it yields results that demonstrate a persistent bias, even if the results are within the prediction interval, it might be beneficial to continue the analysis. If processed samples are somehow related (within one EQAS or PT survey challenge; manufactured from the same formulation, but different batches), then compare the deviations from the line of best fit that has been calculated and drawn from the results of the processed samples with the fresh patient specimen results, either on an aggregate basis or individually. Increasing the number of processed samples, if they are from similar sources, might be of benefit in the evaluation.

6.4 Possible Variations

The following variations to the proposed matrix effect protocols may improve their utility:

(1) Analyze more than 20 fresh patient specimens if the resulting data and plot do not provide adequate information for evaluation.

NOTE: Analysis of more than 20 fresh patient specimens will provide higher sensitivity for identifying the presence of matrix effects. However, the benefits of more samples only increase by the square root of *n* as the number gets larger.

(2) If the scatter around the regression line appears neither proportional nor constant to analyte concentration (i.e., neither variance divided by concentration nor variance itself is constant across analyte concentration), segment the data into several groups of smaller concentration intervals, and perform linear regression analysis within each interval. For example, one would analyze concentration ranges in which the scatter of results from fresh patient specimens about the regression line appears to be approximately constant. A minimum of ten fresh patient specimens that bracket the concentrations of the processed samples should be used within each segment.

References

[1] Fasce CF Jr., Rej R, Copeland WH, Vanderlinde RE. A discussion of enzyme reference materials: applications and specifications. *Clin Chem.* 1973;19:5-9.

[2] ASTM. E456-91a. *Standard Terminology Relating to Quality and Statistics.* Philadelphia: American Society for Testing and Materials; 1991.

[3] ASTM. E177-90a. *Standard Practice for Use of the Terms Precision and Bias in ASTM Test Methods.* Philadelphia: American Society for Testing and Materials; 1990.

[4] Mandel J. *The Statistical Analysis of Experimental Data*. New York: Interscience Publishers; 1964:14-320.

[5] Fisher RA. Mathematical probability in the natural sciences. *Technometrics.* 1959;1:21-29.

[6] Long T. Statistical power in the detection of matrix effects. *Arch Pathol Lab Med.* 1993;117:387-392.

Additional References

Eckfeldt J, Copeland M. *Arch Pathol Lab Med*. 1993;117:381-386.

European Committee for Standardization (CEN). In vitro diagnostic systems — Measurements of quantities in samples of biological origin — Presentation of reference measurement procedures (draft). Ref. No. prEN 12286:1996. Central Secretariat: rue de Stassart 36, B-1050 Brussels, Belgium.

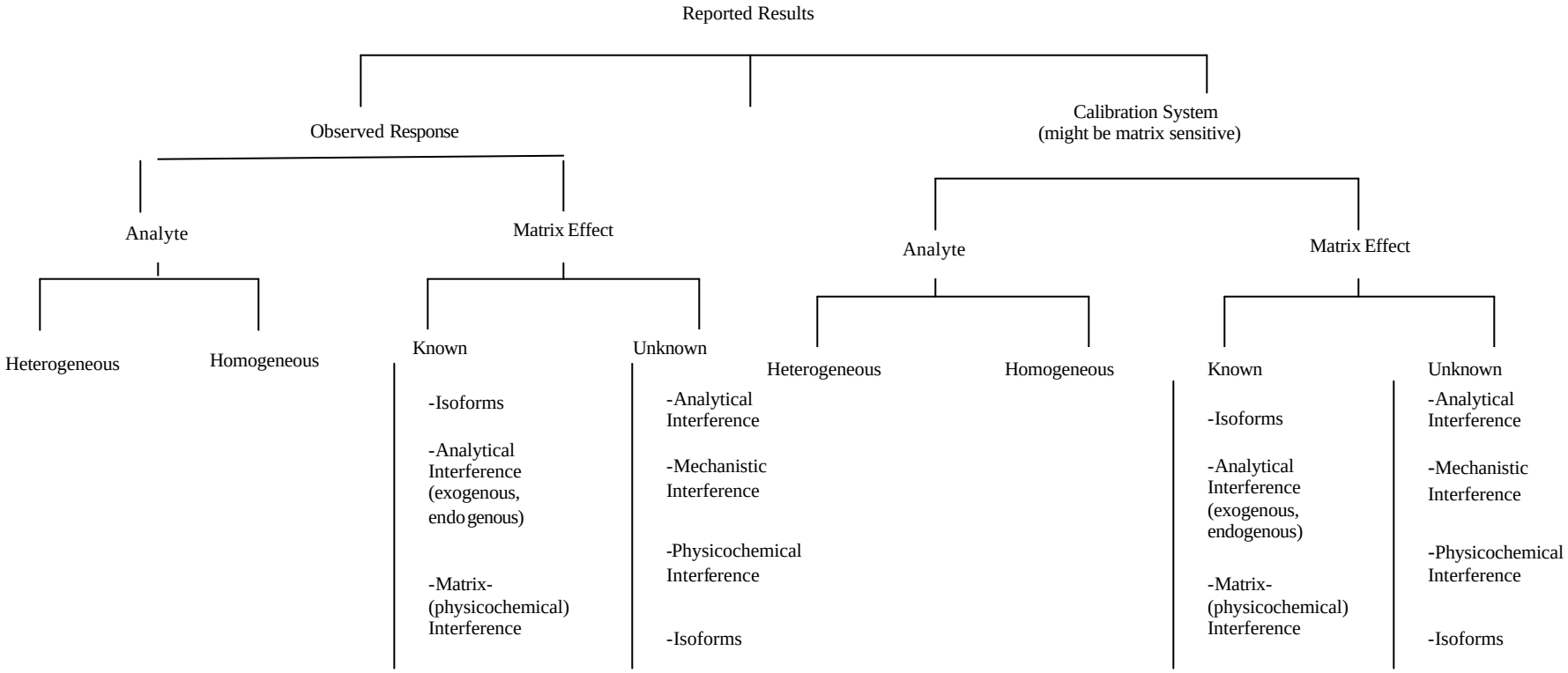
International Organization for Standardization (ISO). *International Vocabulary of Basic and General Terms in Metrology (VIM).* Geneva:1993.

NCCLS. *Apolipoprotein Immunoassays: Development and Recommended Performance Characteristics; Approved Guideline.* NCCLS document I/LA15-A. 1997.

Ross, et al. *Arch Pathol Lab Med.* 1998;122:587-608.

World Health Organization (WHO). *Glossary of Terms for Biological Substances Used for Texts of the Requirements.* Expert Committee on Biological Standardization. WHO unpublished document BS/95.1793. Geneva:1995.

Appendix A. A Hierarchical Diagram of Factors Affecting Reported Results



Appendix B. Data Input Form

	Method Results					
	Comparative			Evaluated		
	#1	#2	#3	#1	#2	#3
Processed Sample #1						
Fresh Patient #1						
Fresh Patient #2						
Fresh Patient #3						
Fresh Patient #4						
Fresh Patient #5						
Processed Sample #2						
Fresh Patient #6						
Fresh Patient #7						
Fresh Patient #8						
Fresh Patient #9						
Fresh Patient #10						
Processed Sample #3						
Fresh Patient #11						
Fresh Patient #12						
Fresh Patient #13						
Fresh Patient #14						
Fresh Patient #15						
Processed Sample #4						
Fresh Patient #16						
Fresh Patient #17						
Fresh Patient #18						
Fresh Patient #19						
Fresh Patient #20						
Processed Sample #5						

Analyte:
Units:

Evaluated method:
Comparative method:

Appendix C. Examples of Completed Analyses

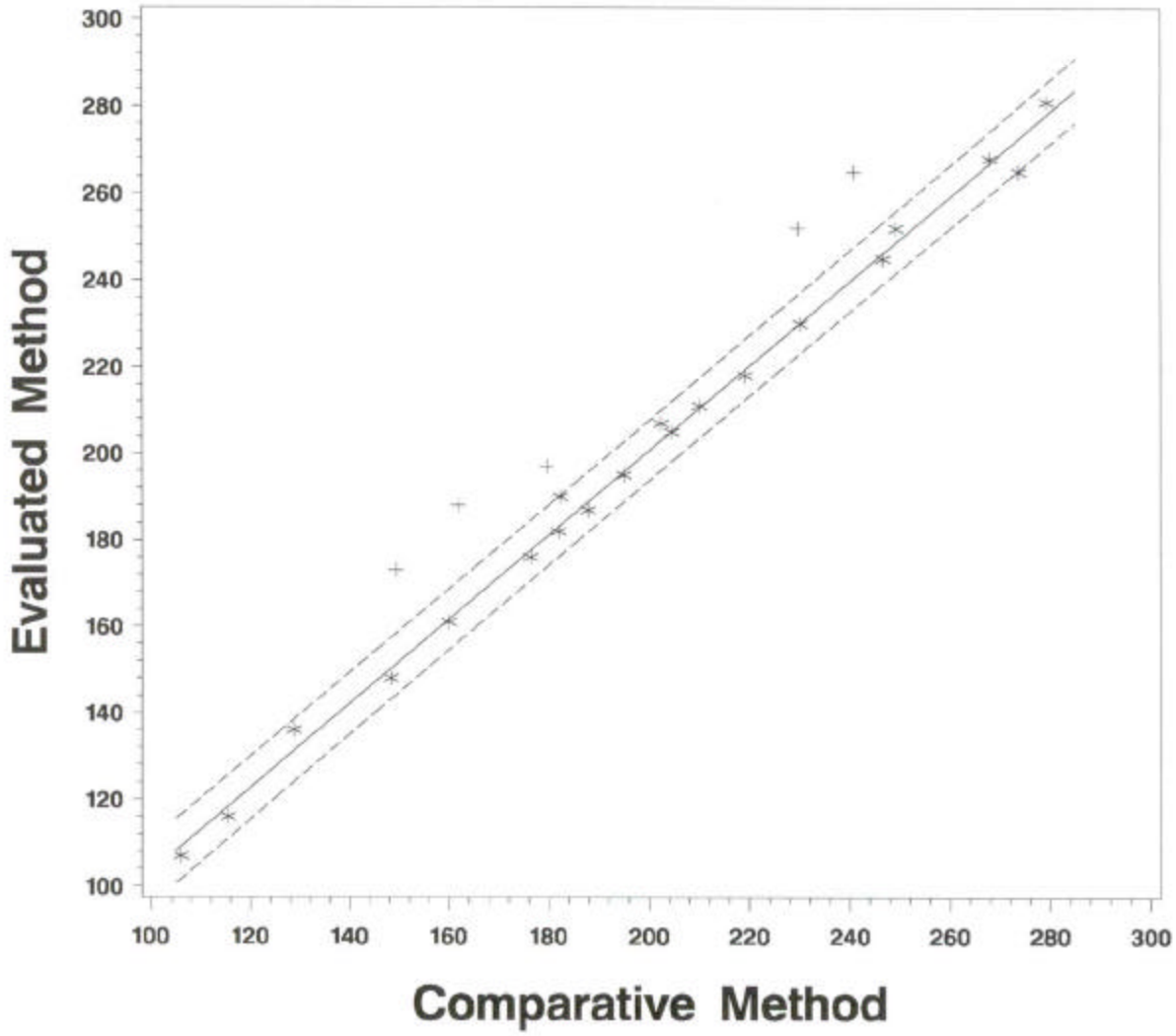
Example 1: Cholesterol; Use of Linear Regression Analysis
See Section 6.3(2)

	Method Results	
	Comparative*	**Evaluated***
Processed Sample #1	**229.6**	**252**
Fresh Patient #1	**246.6**	**245**
Fresh Patie nt #2	**194.9**	**195**
Fresh Patient #3	**267.9**	**268**
Fresh Patient #4	**279.3**	**281**
Fresh Patient #5	**182.3**	**190**
Processed Sample #2	**161.7**	**188**
Fresh Patient #6	**249.2**	**252**
Fresh Patient #7	**115.5**	**116**
Fresh Patient #8	**181.9**	**182**
Fresh Patient #9	**219.1**	**218**
Fresh Patient #10	**128.7**	**136**
Processed Sample #3	**240.8**	**265**
Fresh Patient #11	**148.3**	**148**
Fresh Patient #12	**230.2**	**230**
Fresh Patient #13	**273.7**	**265**
Fresh Patient #14	**159.9**	**161**
Fresh Patient #15	**187.8**	**187**
Processed Sample #4	**149.2**	**173**
Fresh Patient #16	**105.9**	**107**
Fresh Patient #17	**176.4**	**176**
Fresh Patient #18	**202.3**	**207**
Fresh Patient #19	**210.0**	**211**
Fresh Patient #20	**204.4**	**205**
Processed Sample #5	**179.5**	**197**

Analyte: Cholesterol
Units: mg/dL
Evaluated method:
Comparative method:
***mean of three replicates**

Appendix C. (Continued)

Example 1: Sample Plot for Cholesterol



Conclusion: Processed samples exhibit matrix effects that are different from the patient specimens.

NOTE:

***** **Represents fresh patient specimens.**
+ **Represents processed samples.**

Appendix C. (Continued)

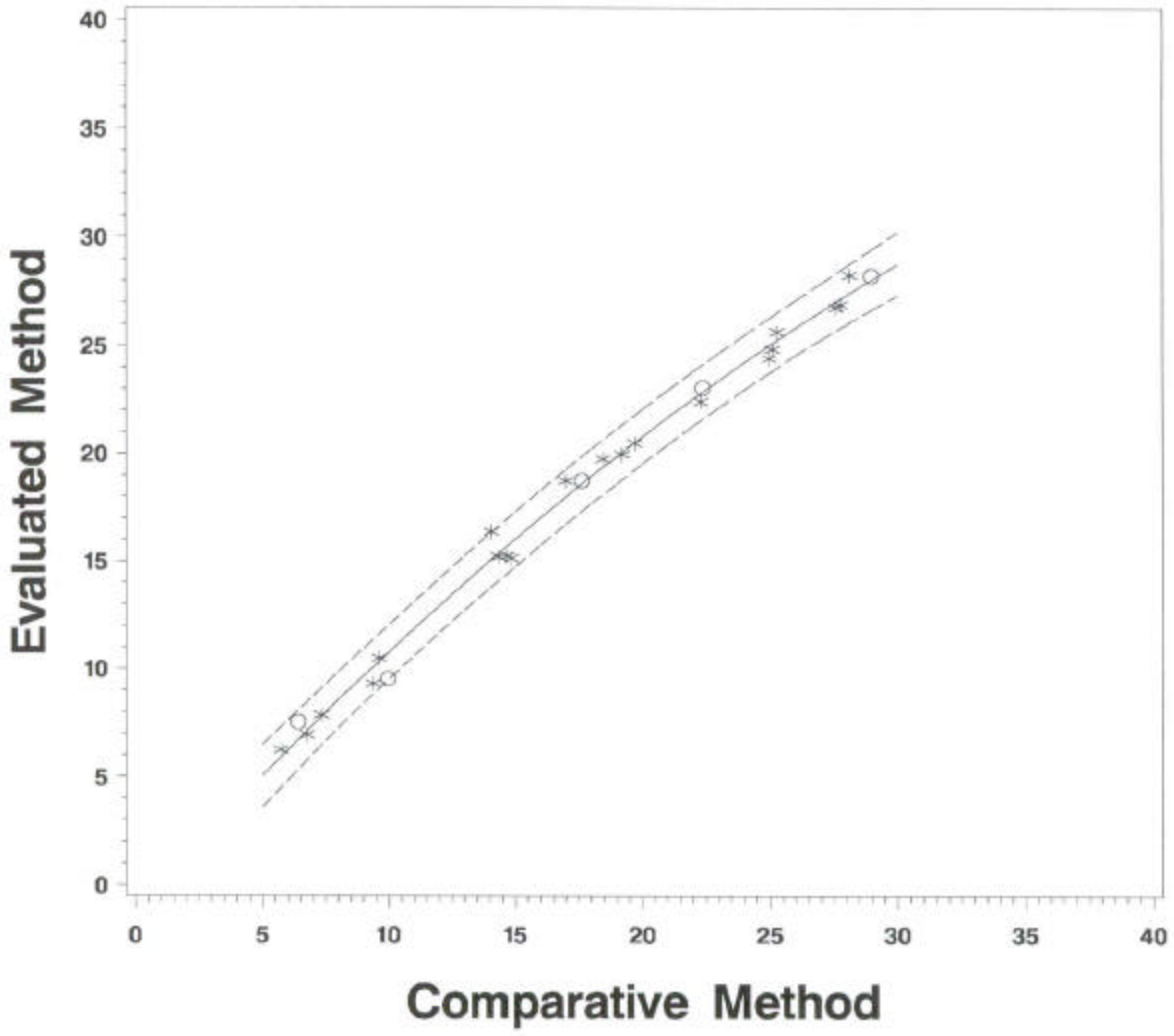
Example 2: Salicylate; Use of Polynomial Regression Analysis
See Section 6.3(2)

	Method Results					
	Comparative			Evaluated		
	#1	#2	#3	#1	#2	#3
Processed Sample	**28.4**	**29.1**	**29.4**	**28.3**	**28.0**	**28.4**
Fresh Patient #1	**27.8**	**27.7**	**27.8**	**27.2**	**26.9**	**26.6**
Fresh Patient #2	**9.6**	**9.6**	**8.9**	**9.6**	**9.1**	**9.2**
Fresh Patient #3	**28.0**	**28.4**	**28.0**	**27.2**	**29.0**	**28.6**
Fresh Patient #4	**7.7**	**6.9**	**7.4**	**8.1**	**7.8**	**7.7**
Fresh Patient #5	**27.4**	**27.8**	**27.6**	**27.7**	**26.4**	**26.2**
Processed Sample #2	**23.0**	**20.9**	**23.2**	**23.4**	**22.8**	**22.9**
Fresh Patient #6	**26.0**	**24.2**	**25.6**	**25.4**	**25.8**	**25.7**
Fresh Patient #7	**25.1**	**25.1**	**25.1**	**24.6**	**25.6**	**24.3**
Fresh Patient #8	**9.6**	**9.4**	**9.8**	**10.6**	**10.7**	**10.2**
Fresh Patient #9	**17.0**	**16.1**	**17.9**	**18.7**	**18.8**	**18.7**
Fresh Patient #10	**25.1**	**24.4**	**25.4**	**24.4**	**24.4**	**24.4**
Processed Sample #3	**18.9**	**16.1**	**17.9**	**18.7**	**18.8**	**18.7**
Fresh Patient #11	**22.4**	**22.4**	**22.1**	**22.5**	**22.4**	**22.3**
Fresh Patient #12	**19.7**	**19.0**	**20.4**	**20.8**	**20.5**	**20.1**
Fresh Patient #13	**20.8**	**17.7**	**19.1**	**20.2**	**20.0**	**19.6**
Fresh Patient #14	**15.3**	**13.3**	**14.4**	**15.2**	**15.1**	**15.3**
Fresh Patient #15	**18.5**	**18.4**	**18.5**	**20.4**	**19.9**	**18.9**
Processed Sample #4	**9.3**	**11.5**	**9.1**	**9.7**	**9.4**	**9.5**
Fresh Patient #16	**14.5**	**15.0**	**14.5**	**14.8**	**15.8**	**15.1**
Fresh Patient #17	**6.9**	**6.6**	**6.8**	**7.1**	**6.9**	**6.8**
Fresh Patient #18	**12.4**	**12.1**	**20.0**	**12.5**	**17.0**	**15.9**
Fresh Patient #19	**5.8**	**5.8**	**5.6**	**6.3**	**6.5**	**5.9**
Fresh Patient #20	**13.1**	**15.8**	**13.2**	**16.2**	**17.0**	**15.9**
Processed Sample #5	**6.5**	**6.8**	**5.9**	**7.8**	**7.6**	**7.2**

Analyte: Salicylates
Units: mg/dL
Evaluated method:
Comparative method:

Appendix C. (Continued)

Example 2: Sample Plot for Salicylate



Conclusion: Processed samples do not exhibit matrix effects.

NOTE:

***** **Represents fresh patient specimens.**
+ **Represents processed samples.**

NCCLS consensus procedures include an appeals process that is described in detail in Section 9.0 of the Administrative Procedures. For further information, contact the Executive Offices or visit our website at www.nccls.org.

Summary of Comments and Subcommittee Responses

EP14-P: *Evaluation of Matrix Effects; Proposed Guideline*

General

1. Matrix effects are a very important item at least in Germany. While in most European countries scientists and lab physicians are well aware of specific matrix effects in processed samples (controls, calibrators), there are experts (professionals) for internal and external quality assurance in Germany who propose that these matrix effects are only indications for a lack of specificity of the methods.

   Therefore, in vitro diagnostic manufacturers in Germany have many troubles, because of the so-called Reference Method Concept. This concept is based on "true values" in controls with reference methods. This causes problems particularly with immune chemical methods because of possibly systematic differences with methods or systems depending values, due to matrix effects. For this reason, this pragmatic NCCLS standard is helpful for the discussion about matrix effects of controls and calibrators.

2. The document proposes to evaluate matrix effects in processed material in the frame of a method comparison between the evaluated and a reference method. Using about 20 different human specimens for the regression analysis a 95% prediction interval is calculated. A matrix effect is seen, if the recovery of the analyte in the processed material is outside of this interval.

   We perform recovery experiments with processed materials having assigned values obtained from appropriate reference methods. The permissible deviation from 100% recovery depends on the imprecision of the evaluation method and the diagnostic and clinical relevance. If a matrix effect exists, then the transferability to an established method has to be demonstrated.

   The experimental requirement in performing both a recovery and a method comparison experiment for testing matrix effects is rather heavy and in many cases not necessary. If one wishes to include a method comparison experiment, then established biometrical considerations should be taken into account.

   - A nonlinear relationship between the two methods is not acceptable; it is caused by a calibration problem in one or both methods. It implies that the concentration scales are not proportional – and I do not know how to interpret this situation.

   - The linear model is only appropriate for those reference methods which produce error-free results. It is advisable to use the structural relationship model for the evaluation of a method comparison, which allows for errors in both variables and is less restrictive with respect to the variances. The corresponding regression procedures are either the DEMING or STANDARDIZED PRINCIPAL COMPONENT or the error-resistant PASSING/BABLOK procedure. This concept is customarily used in Europe and is also known in the U.S.

- The prediction interval decreases with the number of samples, requiring the calculation of sample sizes at predefined error levels. The proposed data structure of three replicate batches would suggest using a more elaborate ANOVA rather than simple regression.

- **The subcommittee, upon review of this substantial input, determined that EP14 need not be changed.**

  The first set of comments reflects the concerns about matrix effects worldwide, not just in Germany. There are many individuals who believe that matrix effects may be due to lack of specificity. That is discussed in the foreword of the document, the third paragraph. So the conclusion is that the subcommittee agrees with the commentor, for the most part. However, to say that lack of specificity is the *only* cause for matrix effects may be an oversimplification of all the potential causes. EP14 is allowing for the possibility of other causes.

  The second set of comments is twofold. The first one is related to the idea of recovery. In paragraph one, the commentor is merely summarizing the approach of EP14. In paragraph two, he is contrasting the one approach from that recommended by EP14. Both require the use of a reference method. The alternative approach then checks to see if the test method is able to recover a result within a certain tolerance (based on the precision of the assay). EP14 recommends the same thing. However, what the alternative approach does not do is to show whether the test method is able to obtain acceptable recovery for patient samples as well; while the EP14 approach does both. What if the alternative approach resulted in poor recovery for all samples: patient samples and artificial samples? It is the difference in recovery between these two types of samples that EP14 is attempting to classify as a matrix effect. The subcommittee is sure that the alternative approach does address this somewhere; it just has not been stated here in the commentor's memo.

  The second set of comments from the commentor is more appropriately directed to EP9 on the determination of bias using a method comparison study. The subcommittee is very familiar with the approach of Passing and Bablok for method comparison. While its major benefit is to enable a determination of the slope of a comparison line between two sets of data in a way that makes no assumptions about the characteristics of the data, it requires extensive amount of computer power to perform—something that may inhibit its use in a routine hospital laboratory. So recommendations or references to this technique should be referred to EP9 for consideration there.

  Regarding nonlinearity in the data, we agree with the commentor, that ideally there should be no nonlinearity. The purpose of including nonlinear regression is to prevent the inflation of the statistic that is used to calculate the tolerance limits for the patient sample results. If the data are such that a nonlinear regression provides a lower sum of squares of residuals, then that is the preferred analysis to identify the relation of the response of the test method to the "reference method." This will provide an appropriate assessment of the tolerance limits, and will result in improved sensitivity to possible matrix effects. While each assay may have been independently determined to be "linear," when their results are compared, the comparison process might accentuate the different linearization functions used by each assay. To ignore nonlinear relations by forcing them to be linear merely causes an inflated standard error and inflated tolerance limits. This could reduce the sensitivity of the process to detection of matrix effects.

3. This document is in line with NCCLS standards. It is mostly applicable to manufacturers, regulators, and proficiency programs and is of some use to clinical laboratories.

- **The subcommittee appreciates this input. The objective of the subcommittee was to provide an NCCLS guideline that had broad applicability.**

Foreword

4. Text within the foreword stating no generally accepted guidelines for identifying or quantitating matrix is erroneous. The paper by Eckfeldt and Copeland, *Arch Pathol Lab Med* 1993;381-386 has been used for several years to identify the presence of a matrix bias and is the basis for the protocol in this document. Two papers have been published which provide methodology to quantitate matrix interference: Ross, et al. *Arch Pathol Lab Med* 1993;117:393-400 and Ross et al. *Arch Pathol Lab Med* 1998;122:587-608.

- **The subcommittee agrees. The following terminology has been added: "…analytical methods, when this work was first proposed there were no generally…"**

5. The foreword should mention that information from EP9 is needed for data evaluation in the EP14 protocol.

- **The subcommittee disagrees; the approach to evaluate patient specimens is similar, yet important differences exist, specifically: a) the number of samples recommended for the evaluation (20 for EP14); and b) the concentration range of patient samples recommended in EP14 is best determined by the concentrations of the processed samples – full reportable range measurements are not required.**

Section 4

6. Regarding the term "analytical interference." This is a confusing definition. In the mechanistic interference and physicochemical interference sections it is not clear what is meant by "the population mean of the test results" and what relationship this has to an analytical interference. In the physicochemical section, the words "an environmental or structural difference from that of the patient specimens" is difficult to understand. The word "structural" can relate to different iso forms, changes in tertiary structure of proteins, altered haptenic groups or the presence of antigenic components or molecular forms not found in native clinical specimens. What does the word "environmental" relate to? The words in the definition are cryptic and convey little meaning. Consider limiting the definition to the first two sentences. Include the explanatory detail in the text section 5 (see #8 below).

- **The term "population" has been deleted from this section. The examples provide reasonable illustration of what the subcommittee intended to convey. It is often useful to include definitions from other standards and guidelines (as we do here). Without alternative wording that might have been considered, the definitions stand, as agreed to by the subcommittee.**

7. In the definition of average bias; line 2 insert "or comparative" between reference and method.

- **The subcommittee agrees, and the wording has been changed accordingly.**

8. In the definition of calibration; line 8 "or values…" is confusing. The material in "NOTES" is unintelligible to me.

- **These definitions are all referenced. Each is understood by the group that established it. The last definition, which was established by metrologists, is considered to be the most precise by**

that group (VIM is an accepted and widely recognized standard); it has been provided here to demonstrate that all these definitions apply for use with EP14.

Section 5

9. First paragraph: temperature is not usually considered a "matrix effect" but rather a variable to be controlled in a reaction.

- **Temperature is an environmental condition that can affect measurement. As explained in the Foreword, "matrix effect" is a pragmatic definition that might call for further investigation, such as how carefully temperature must be controlled.**

10. First paragraph: in addition to activity and concentration, other parameters can be measured, for example cell count or number of organisms. Can the document address these kinds of measurements?

- **The subcommittee did not have adequate information or data to address these types of measurements; these may be addressed in future revisions of EP14.**

11. Second paragraph: the second sentence blurs the distinction between a method nonspecificity interference and a matrix interference. The confusion identified in point (3) above relates to the same blurring of these two issues. Analytical interference is a general term that includes a variety of sources and/or types of method nonspecificity for the analyte of interest. In most cases the term analytical interference (perhaps method interference could be used) is used to refer to substances in the native clinical specimen that can cause an erroneous measurement of the analyte of interest. A matrix interference is a type of analytical interference which represents the special case when the interference is introduced to the native specimen matrix by some process that manipulates or changes the native clinical material. The matrix change is generally a byproduct of unavoidable practical constraints of manufacturing, distribution, analyte content requirements, etc. The net effect of a matrix interference is that a processed specimen, that otherwise might be expected to behave the same as a native clinical specimen, has an altered performance or response with the measurement system.

    I realize these definitions of terms and concepts are at the core of this document and are difficult to define in a way that satisfies all concerns. However as presented in this draft the term matrix interference is confusing and does not assist the profession to define and understand this phenomenon.

- **We refer the commentor to the Foreword and Introduction. The subcommittee struggled with the same issues and worked hard to promote a clear and relevant discussion and definition. In light of no suggested alternative, the wording stands.**

12. Third paragraph: lines 5-7. The same confusion of terms mentioned above carries into this section. The primary contributor to scatter is the imprecision and analytical nonspecificity of the field method vs. the comparative method. The replication is not adequate to reduce the measurement imprecision to zero. The method nonspecificity in field or comparative methods can be due to various types of analytical interferences. However the essential premise of this protocol is that the native clinical specimens are assumed not to have a "matrix interference" component. Only with this assumption can the protocol identify the presence of a matrix interference in a processed material under evaluation. Thus the distinction in terms detailed in (8) above.

- **The subcommittee disagrees (although not entirely) on two points: a) the imprecision is significantly reduced by measuring replicates on all patient samples. Obviously, it can never be brought down to zero, but variances from 20 samples can be pooled to reduce the variation due**

to repeatability; and b) the use of 20 specimens is to obtain an "average" estimate of the (relative) nonspecificity of the method. The number of total measurements was chosen for practical and statistical reasons.

13. Last paragraph, last sentence: change "might" to "will."

- **The subcommittee disagrees with this editorial change. It cannot be predicted how such studies are designed and conducted, nor the variables that are involved. The statement provides more flexibility as written.**

Section 6.1

14. Second bullet, "Note" line 1. The phrase "comparative method should be free of matrix effects" is incorrect. The matrix effect is a joint property of a specific method/material combination. Some methods may be prone to occurrence of matrix effects with various materials but the term matrix effect is not an attribute of a method. A better phrase would be "comparative method should be free of matrix related interferences."

- **The phraseology has been edited as follows: "Although ideally, the comparative…". A point of significance is that an analytical method must also include a definition of the samples for which it is intended. Therefore, no further changes are needed.**

15. What about SRMs (Standard Reference Materials) 256 and 965?

- **These materials are only useful if a reference method is used as the comparison method, and the use of high-quality materials for standardization should be included in the method's operating procedure.**

Section 6.2

16. Numbers (2) and (3). Reconsider the recommendation "preferably with separate calibrations." The rationale to not recalibrate, when the method calibration is stable, is that recalibration will introduce a source of imprecision that can otherwise be avoided.

- **The strategy is to increase the normally encountered variabilities to make it more difficult to "declare" a matrix effect (the lines of acceptance are wider). We wanted to avoid an apparent matrix effect that could be caused by a "low" or "high" calibration.**

17. Number (4): consider changing "enzyme" in the last line to "molecular" to confer more generality to the statement.

- **The statement is an example and a general fact as written.**

Section 6.3

18. Bullet 2, second line; insert "from patient specimens" after results. This will ensure clarity about which data to use in this step.

- **This section has been edited as suggested.**

19. Section 6.3(5) end of first paragraph: there is a reference to Section 6.4(6) which does not exist.

- **This cross-reference has been changed to read "6.4(2)."**

Appendix A

20. Appendix A is confusing and adds no value to the document. I did not find where in the text this appendix is referred. If this diagram is retained then the concepts must be clarified as discussed in (3), (8) and (9) above.

- **A statement has been added to the end of paragraph 5 in the Foreword to refer to Appendix A.**

Appendix C

21. Page 16. Appendix C, example 2. The figure presentation needs to be improved to journal publication quality. The symbols for processed materials cannot be distinguished from patient specimens. Page 14, example 1 would benefit from higher quality as well although the plot symbols can be distinguished.

- **The graphic has been edited for clarity.**

Summary of Delegate Voting Comments and Subcommittee Responses

EP14-A: *Evaluation of Matrix Effects; Approved Guideline*

General

1. Pages 7, 14, and 15: Based on the table of page 14, the plot for example 1 has the "*x*" and "*y*" variables switched. Item (1) of section 6.3 clearly states that the plot should be made with "the evaluated method results on the *y* axis and the comparative method on the *x* axis". I did not check to see if the same error is repeated in example 2.

- **New plots have been added.**

2. Pages 8 and 14: It is a fundamental statistical fact that the 95% prediction limits as denoted in equation (2) apply only to a *single* prediction made from the calculated regression line. In the examples, you are evaluating the prediction intervals for 5 different values of *X*, and the confidence level should be adjusted accordingly. As presently stated, the matrix effect evaluation has an overall confidence level of 75%. There is a very simple solution, which is to divide the error level α by the number of processed samples in the experiment. In the notation of equation (2), this means using t (0.995, *n-2*), where 0.995 is $\alpha/(2*5)$.

- **The subcommittee disagrees. The intent is to evaluate *individual* processed sample biases with the correlation confidence interval of a population of 20 (plus) patient samples. Processed samples which are interrelated are addressed in Section 6.3(6).**

3. Page 8: In step (4), we want prediction intervals for the processed samples, not the fresh patient specimens. While this is not an important distinction if the only goal is to produce a plot with the confidence levels drawn in, it fails to instruct users on how to construct the actual prediction intervals of interest. Equation (2) could be more generally stated by replacing the final term with something like

$$\frac{(\overline{X}^{*} - \overline{\overline{X}})^2}{\sum(\overline{X}_i - \overline{\overline{X}})^3}$$

where X^* is any value of X for which we would like to obtain prediction interval, $\overline{\overline{X}}$ is the overall mean of the comparative method results for the fresh patient specimens, and the $\overline{X}_i$'s are the comparative method results for the fresh patient samples. Similarly, the definition for $\overline{Y}_{pred}$ should be the predicted value at any value of X^*, based on the estimated regression curve, and the definition of the regression standard error should be clarified by using more standard notation and avoiding potential confusion between Y_{pred} and $\overline{Y}_{pred}$.

To generate the plots as suggested in the appendix, the user could be instructed to replace X^* with any grid of values that span the *x*-axis. To calculate the prediction intervals for the processed samples, the user could be instructed to replace X^* with the comparative method mean for each processed sample.

- **The subcommittee disagrees. The standard for comparison is the scatter of patient samples around the line of best fit. The scatter is a good estimate of the relative nonspecificity of the two methods that is presumably due to matrix effects (the confounding influence of variability of**

reproducibility of the measurement is reduced by replicate measurements). The confidence interval calculated from the measurements of the processed sample (as an individual material) represents the reproducibility of the measurement, i.e., precision. It is not appropriate to compare the variability of these two independent parameters.

4. Pages 8 and 9: There is no guidance provided if some of the processed samples fall within the prediction intervals and some do not. Is the user to conclude that there is a matrix effect at some of the comparative method levels while not at others? It would be useful to see some discussion of this potential outcome.

- **The prime evaluation technique is for individual processed samples. For samples that are inter-related, i.e., identical matrix, see 6.3(6).**

Section 6.1

5. Care must be taken to ensure that the analyte range of processed samples, and thus of the patient specimens, is adequate for reliable estimation of regression statistics. It is conceivable that a particular set of PT samples might fall within a relatively tight range. In that case it would be desirable for the patient specimens to span a wider range. Refer users to the "Test for Adequate Range of X" in EP9-A (Section 4.5).

- **This is adequately covered in Section 6.1, Bullet 4.**

Section 6.2(2) and 6.2(3)

6. As a minor upgrade, suggest that the patient and processed samples be run in a different random order for each of the three replicate batches.

- **It is the opinion of the subcommittee that this item is adequately covered in 6.2(2). This comment will be considered in the next revision.**

Section 6.3(2)

7. I have the same concern about nonlinearity between methods as expressed in the first item of the comment section. While I understand the subcommittee's wish to provide an analysis approach for such a case, I think there should be a warning in the data analysis section that this is fairly unusual behavior. Users should be wary of such cases and verify the relationship by a more extensive method comparison, not just go ahead blindly with the polynomial analysis.

- **The subcommittee disagrees. Running more samples will not make a method linear. As mentioned in the document, if the method "appears" nonlinear, but the curvature is a function of the population of patient samples used, it will be more difficult to state that a matrix effect is present. Therefore, the user would logically expand the test to include more samples if that is necessary. The intent of the evaluation is met with the protocol recommended.**

Appendix C, Sample 2

8. Is there a typo in replicate 3 for the comparative method with fresh patient #18? The discrepancy between these 3 values is considerably larger than that seen for any other samples by either method.

- **This has been corrected; two new plots have been added.**

Related NCCLS Publications*

EP5-A **Evaluation of Precision Performance of Clinical Chemistry Devices; Approved Guideline (1999).** This document provides guidance for designing an experiment to evaluate the precision performance of clinical chemistry devices; recommendations on comparing the resulting precision estimates with manufacturer's precision performance claims and determining when such comparisons are valid, as well as manufacturer's guidelines for establishing claims.

EP6-P **Evaluation of the Portable Range (Linearity) of Quantitative Analytical Methods; Proposed Guideline (1986).** This document contains a method for evaluating whether an instrument or quantitative analytical method on the basis of the manufacturer's linearity claim; and offers guidelines for manufacturer's use when stating a claim of an assay's linear range.

EP7-P **Interference Testing in Clinical Chemistry; Proposed Guideline (1986).** This document provides background information and procedures for characterizing the effects of interfering substances on test results.

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.

I/LA15-A **Apolipoprotein Immunoassays: Development and Recommended Performance Characteristics; Approved Guideline (1997).** This guideline describes the characterization and preparation of immunogens, antibodies, samples, and methods, and provides guidance for immuno-chemical testing of apolipoproteins.

M29-A **Protection of Laboratory Workers from Instrument Biohazards and Infectious Disease Transmitted by Blood, Body Fluids, and Tissue; Approved Guideline (1997).** 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 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).

NRSCL13-A **The Reference System for the Clinical Laboratory: Criteria for Development and Credentialing of Methods and Materials for Harmonization of Results; Approved Guideline (2000).** This document describes procedures for developing and evaluating definitive methods, reference methods, designated comparison methods, and reference materials to provide a harmonized clinical measurement system.

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

NOTES

NOTES

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 1-56238-434-1