

Workshop Report

Workshop on Bioanalytical Methods Validation for Macromolecules: Summary Report

Authors: Krys J. Miller,¹ Ronald R. Bowsher,² Abbie Celniker,³ Jacqueline Gibbons,⁴ Shalini Gupta,⁵ Jean W. Lee,⁶ Steve J. Swanson,⁷ Wendell C. Smith,⁸ and Russell S. Weiner⁹

Contributors: Daan J. A. Crommelin,¹⁰ Ira Das,¹¹ Binodh S. DeSilva,¹² Robert F. Dillard,¹³ Michael Geier,¹⁴ Han Gunn,¹⁵ Masood N. Khan,¹⁶ Dean W. Knuth,¹⁷ Michael Kunitani,¹⁸ Gerald D. Nordblom,¹⁹ Rene J. A. Paulussen,²⁰ Jeffrey M. Sailstad,²¹ Richard L. Tacey,²² and Ann Watson²³

Received March 7, 2001; accepted May 10, 2001

KEY WORDS: bioanalytical methods validation; macromolecules; ligand-binding assay; immunoassay; ELISA; EIA; RIA; anti-drug; bioassay; biomarker; total error; assay; quality control; acceptance criteria; specificity; accuracy; precision; recovery; calibration; drug development; neutralizing antibody.

INTRODUCTION

This report summarizes the outcome of a workshop on "Bioanalytical Methods Validation for Macromolecules"²⁴ that was held in March 2000 in the Washington, DC area. The workshop was principally sponsored by the American Asso-

ciation of Pharmaceutical Scientists (AAPS)²⁵ and had the following goals:

1. To determine industry and regulatory standards established for bioanalytical method validation in support of the estimation and characterization of macromolecules in the pre-clinical and clinical stages of drug development.
2. To evaluate special validation considerations for quantitative, macromolecule-detecting technologies that have emerged since 1990 including immunoassays, cell-based assays, antibody titers, and automation in the laboratory.
3. To address the strengths/limitations and advantages/disadvantages of assay-customized approaches to validation that focus on assay parameters specific to the intended use of the assay.
4. To develop a 2000 workshop report regarding appropriate bioanalytical validation criteria and standardization of terminology for the above to be used by regulatory agencies to draft new guidelines for the bioanalytical validation criteria for macromolecules quantitation.

The purpose of this report is to summarize the major issues and recommendations from the workshop, thereby providing guiding principles for the validation of bioanalytical methods used to support the preclinical and clinical stages of macromolecular drug development. The workshop was organized into multiple sessions that addressed discrete topics germane to methods validation. At the close of the workshop, each session's chairperson was asked to prepare a summary of

¹ Pharmacokinetics and Drug Metabolism at Amgen, Inc., One Amgen Center Drive 1-1-B, Thousand Oaks, CA 91320. To whom correspondence should be addressed. (e-mail: krysm@amgen.com)

² Eli Lilly and Co., Indianapolis, IN 46285.

³ Genetics Institute/Wyeth Ayerst Research, Andover, MA 01810.

⁴ Jacqueline Gibbons Consulting, Piedmont, CA 94611.

⁵ Clinical Immunology at Amgen, Inc., Thousand Oaks, CA 91320.

⁶ Jean Lee, MDS Pharma Services, Lincoln, NE 58502.

⁷ Clinical Immunology at Amgen, Inc., Thousand Oaks, CA 91320.

⁸ Eli Lilly and Co., Indianapolis, IN 46285.

⁹ Bristol-Myers Squibb, New Brunswick, NJ 08903-0191.

¹⁰ University of Utrecht, Utrecht, NET, Netherlands.

¹¹ Washington University, St. Louis, MO 63130.

¹² Procter & Gamble Pharmaceuticals, Mason, OH 45040.

¹³ Pharmacia Corporation, Skokie, IL 60077.

¹⁴ Roche Molecular Systems, Pleasanton, CA 94588-2722

¹⁵ Pharmacokinetics and Drug Metabolism at Amgen, Inc., Thousand Oaks, CA 91320.

¹⁶ MedImmune, Inc., Frederick, MD 21703-8619.

¹⁷ Pharmacia & Upjohn, Inc., Kalamazoo, MI 49007.

¹⁸ Corte Madera, CA 94925.

¹⁹ Warner-Lambert/Parke-Davis, Ann Arbor, MI 48105.

²⁰ Phoenix Int'l Life Sciences, St. Laurent, Quebec, Canada, H4R 2N6.

²¹ Glaxo Wellcome, Inc., Res. Triangle Park, NC 27709.

²² PPD Development, Richmond, VA 23230.

²³ Pharmacokinetics and Drug Metabolism at Amgen, Inc., Thousand Oaks, CA 91320.

²⁴ "Macromolecule" is defined as a molecule (generally a biopolymer) having a mass greater than 1,000 daltons.

²⁵ This report represents the composite opinion of the participants of the workshop and not necessary the view or policy of AAPS or other organizations.

the key points discussed with the goal of capturing the "spirit" of the session to the best of their ability. This report represents a compilation of the summaries prepared by the chairpersons. It is noteworthy that scientists in this area currently have divergent views regarding validation of assays for macromolecules. Since the recommendations in this report do not necessarily reflect a concordance of viewpoints, each investigator is obliged to formulate an interpretation of the principles described in this report and to justify the interpretation when applying principles to practice.

The organization of this report is similar to that of the workshop. Thus, after a brief overview of historical perspectives on macromolecular bioanalytical method validation, 5 topics are sequentially addressed: 1) Standard immunoassays; 2) Anti-drug antibody assays; 3) Bioassays (cell-based and activity-based); 4) Biomarker assays; and 5) Validation and acceptance criteria. Finally, this report closes with a glossary of terms relating to the bioanalysis of macromolecules.

II. HISTORICAL PERSPECTIVES

In 1990 a conference was held to discuss bioanalytical method validation, the results of which were published in a 1992 conference report by Shah *et al.* (1). In January 2000 a similar conference was held to revisit the issues discussed in 1990 and to update the practices and validation recommendations to better reflect the state of the art in the current technology (2). The 1990 and January 2000 conferences discussed ligand-binding assays and/or microbiological assays, but the main focus was bioanalytical methods that quantify low molecular weight molecules using chromatographic and mass spectrometric techniques. Recommendations outlined in the 1992 report and subsequent guidelines have been implemented for the past eight years for assaying both small molecules and macromolecules. Although the need for a broader discussion related to the non-chromatographic bioanalytical methods was evident, it was in 1998 when the specific challenges related to the GLP compliant validation of ligand-binding assays were discussed in a roundtable at the Annual meeting of the AAPS in San Francisco. This roundtable discussion subsequently resulted in the development of a position paper (3).

As discussed in the position paper (3), there are inherent differences between bioanalytical methods applied to pre-clinical and clinical development of small molecules versus macromolecules. Thus, whereas small molecules are commonly analyzed in chromatographic assays (e.g., HPLC), macromolecules are principally measured using immunoassay and bioassay techniques.²⁶ Detection in chromatographic assays is based on the physicochemical properties of the small molecule analyte, but detection of macromolecules is normally accomplished using reagents that are derived from living systems (e.g., cell lines), with the attendant variation that arises from biological products. Measurement of small molecules in *ex vivo* samples is routinely preceded by physical isolation from molecular species that could interfere with analyte detection. However, because macromolecules are

so highly potent, their molar concentrations in *ex vivo* samples are typically too low to support molecular isolation. In addition, current separation technologies are inadequate (e.g., solid phase or liquid/liquid extraction) or impractical (e.g., capillary electrophoresis or size exclusion chromatography) for isolation of macromolecules. Therefore, the detection of a macromolecular analyte generally occurs in a complex biological milieu, a process that is highly dependent on the integrity of reagents used in the detection system. As a consequence of these and other factors, macromolecular bioanalytical methods tend to have poorer batch-to-batch reproducibility than bioanalytical methods involving small molecules. Finally, statistical considerations underlying the calibration of bioanalytical methods for small molecules may be inappropriate for calibrating macromolecular bioanalytical methods (3), and guidelines issued for small molecules may therefore lead to imprecise results when applied to macromolecules.

In conclusion, in 1999 it was recognized that the draft guidelines issued for the analysis of small molecule pharmaceuticals are largely inappropriate for the bioanalysis of macromolecules. Thus the March 2000 workshop on "Bioanalytical Methods Validation for Macromolecules" was convened to address the potential issuance of new guidelines suitable for the analysis of macromolecules from biological matrices.

III. LIGAND-BINDING ASSAYS

The term "ligand-binding assay" broadly refers to methods that depend on specific binding of an analyte to a biomolecule. Typically these are reversible binding events governed by the laws of mass action. Since the vast majority of bioanalytical methods currently being used for macromolecules are based on ligand-binding assays, this assay format was the principle focus of the workshop. Included in the family of ligand-binding assays are standard immunoassays, assays to measure anti-drug antibodies, bioassays and biomarker assays.

Ligand-binding assays often involve reagents derived from biological sources and may be subject to inherent variation. Therefore care should be taken so that bioanalytical method development does not proceed into advanced stages while performing in a fashion that is incompatible with a validated bioanalytical method. Rather than being undertaken subsequent to full development, validation of bioanalytical method performance should be integrated into the development of the bioanalytical method from the earliest stages.

In the interest of promoting harmonization, an attempt has been made to conserve the general framework for methods validation outlined in the 1990 Conference Report (1). Accordingly, the present report divides the process of method validation into 3 discrete phases: 1) Bioanalytical Method Establishment (e.g., Reference Standard Preparation); 2) Pre-Study Validation for Bioanalytical Method Establishment; and 3) In-Study Validation (Table I). In the sections that follow, guidance is provided for application of the three-phase procedure when validating each of the major ligand-binding bioanalytical methods.

A. Standard Immunoassays

The ligand-binding assay in predominate usage today is the immunoassay, a bioanalytical method in which antibodies

²⁶ Chromatographic assays are commonly applied to the analysis of protein drugs, but this almost exclusively occurs during the characterization and release of drug products, areas that are beyond the scope of the present manuscript.

Table I. Validation Procedures Outlined in the 1990 Conference Report on Analytical Methods Validation (1)

3 Phases of Bioanalytical Methods Validation
1. Bioanalytical Method Establishment
2. Pre-Study Validation
a. Specificity
b. Calibration Model
c. Precision, Accuracy, Recovery
d. Quality Control Samples
e. Stability
f. Acceptance Criteria
3. In-Study Validation

are the principle reagents. Within the family of immunoassays is the radioimmunoassay (RIA), a pioneer technique for measuring macromolecules that is still used in some biopharmaceutical settings. A more widely used immunoassay technique is the enzyme linked immunosorbent assay (ELISA or EIA), the bioanalytical method that served as the reference method for discussions at the workshop (Fig. 1).

The foregoing discussion of standard immunoassays addresses validation of methods that derive quantitative measurements of analytes in biological matrices such as blood, plasma, serum, or urine, with recommendations pertaining to both clinical and preclinical studies. It should be understood that immunoassays may be used to derive semi-quantitative measurements, but these applications entail special considerations that are covered in separate sections of the report: Anti-Drug Antibody Assays and Biomarker Assays.

1. Bioanalytical Method Establishment

Bioanalytical methods invariably include the use of calibration standards and quality control (QC) samples in which a reference standard is spiked into blank analytical samples. Thus the quality of the reference standard can have a large impact on the integrity of the bioanalytical data. Considering that the analytical reference standard is pivotal to the success of the immunoassay in deriving accurate measurements, special care should be taken when selecting the standard. While

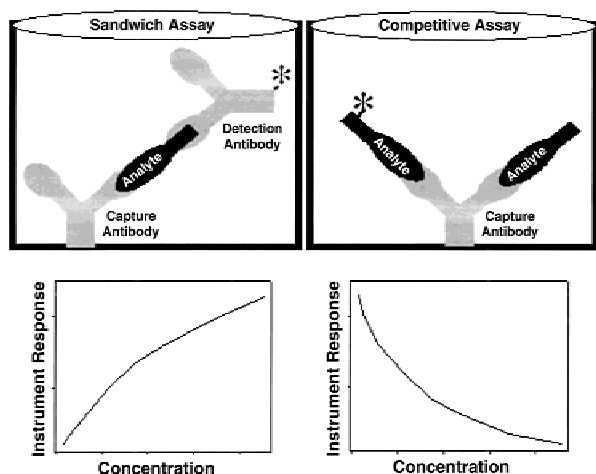


Fig. 1. Common formats for enzyme linked immunosorbent assays (ELISAs or EIAs).

it is generally recommended for bioanalytical methods involving small molecules that the reference sample be procured from a certified reference standard (e.g., USP compendial standards), a commercial supplier, or custom-synthesis by a bioanalytical laboratory, such sources are generally not viable for macromolecular biopharmaceuticals. Rather, the innovator company is typically the most reliable supplier of an authentic reference sample for macromolecular drugs. Therefore, in most circumstances, an authentic reference sample for an immunoassay method should be procured from the innovator company and should correspond to a specified manufactured lot of the drug product.

The concept has been put forward of comparing all reference standards to a Master Standard, a synthetic batch for which identity and purity are clearly established. However, it should be recognized that in clinical and preclinical studies, the important issue is not always comparison to a Master Standard, but the equivalency of the reference standard to the material used in the clinical study. It is therefore recommended that tests for bioanalytical equivalence of the reference standard be tailored to the intended application of the bioanalytical method.

For Method Establishment, documentation of data relating to the authentic reference standard should include a record of the lot number(s), certificates of analysis and stability (when available) and information regarding the identity and purity of the reference standard.

In addition to procuring an appropriate reference standard, Bioanalytical Method Establishment for Standard Immunoassays should focus on selecting conditions for achieving the desired method performance in terms of specificity, accuracy, dilutional linearity, etc. The selection of an appropriate diluent and the minimal required dilution need to be addressed during the development of a bioanalytical method because of the heterogeneous nature of the samples typically analyzed by immunoassays. The most important reagent in an immunoassay is the antibody (or antibody pairs) used to measure the analyte. Antibody reagents must be shown to bind the analyte(s) with suitable specificity, and where the immunoassay involves a sandwich ELISA format, an antibody pair must be chosen which binds the analyte in a complimentary fashion.

2. Pre-Study Validation

A. Specificity. Specificity has been previously defined as the ability of a bioanalytical method to differentiate and quantitate the analyte in the presence of other constituents in the sample (4). However, macromolecular analytes commonly hold structural elements in common with endogenous molecules, and it may therefore be impossible to develop an immunoassay that completely differentiates the analyte from other sample constituents. Rather, investigation of specificity during Pre-Study Validation should focus on reliable quantitation of the analyte against a background of interference from endogenous matrix components. In confirming that the immunoassay method is adequately specific, measurements should be assessed when the analyte is spiked at concentrations throughout the assay range, with an emphasis on spikes near the limit of quantitation, and these measurements should be made in the appropriate biological matrix (e.g., serum) from a number of representative subjects.

B. Assay Calibration. The calibration model (standard curve) is a critical aspect of bioanalytical method performance. The concentration-response curve for an immunoassay typically has no less than 8 non-zero calibrators in duplicate, and generally includes calibrators outside of the validated range of the assay to serve as anchor points that facilitate curve-fitting. When establishing the mathematical model that best describes the standard curve, it should be recognized that the concentration-response curve for a typical ligand-binding assay is seldom linear throughout, but rather sigmoidal, and commonly best described by a logistic function. To limit bias during the regression analysis, in some cases it is advisable to use weighting schemes (such as $1/y$, $1/y^2$, or computed values based on smoothed variance functions (3)). It is beyond the scope of this report to address tests for evaluating the aptness of the model for the particular data at hand, yet it deserves to be noted that the correlation coefficient (R) and the coefficient of determination (R^2) are unreliable for evaluating curve-fits for nonlinear calibration models.²⁷

C. Precision, Accuracy and Recovery. Quality Control (QC) samples (see III.A.2.D. section) should be used to evaluate whether the bioanalytical method has acceptable precision and accuracy. The criteria for precision and accuracy during Pre-Study Validation of an immunoassay method are described in the Validation and Acceptance Criteria section of this report.

Recovery corresponds to the instrument response with an amount of the analyte spiked into and recovered from the sample matrix, compared to the instrument response for the pure reference standard. Recovery is an important indicator of method suitability for chromatographic assays in which procedures used to isolate the analyte from interfering matrix components can result in significant loss of the analyte. However most immunoassays do not involve prior sample extraction, and recovery determinations are therefore of lesser importance. Moreover, in a typical immunoassay, components of the sample matrix can indirectly affect the generation of signal by affecting the binding of the detector antibody. Therefore, instrument response (a downstream result of signal generation) is commonly affected by the sample matrix, making it difficult to assign appropriate conditions for measuring the reference standard. Accordingly, Pre-Study Validation for standard immunoassays generally need not include assessments of analyte recovery. In the limited number of cases in which immunoassay techniques incorporate prior sample extraction, as long as recovery is acceptably reproducible, the extent of analyte recovery (e.g., X% recovery) is only of relevance if it becomes inconsistent across the concentration range.

Although every attempt is made during development of an immunoassay method to reduce matrix interference, it should be understood that immunoassays commonly suffer from nonspecific instrument responses due to reagent cross-reactivity with matrix components. Consequentially it is advisable to evaluate dilutional linearity using serially diluted spiked samples.

D. Quality Control (QC) Samples. Immunoassay bioanalytical methods should include Quality Control (QC) samples, samples having known concentrations of the analyte that are essentially treated as unknowns in the assay. QC samples are generally prepared at concentrations that fall within the linear range of the calibration curve and are prepared in a medium that emulates the matrix of the study samples. During Pre-Study Validation, QC samples are primarily used to estimate assay parameters (i.e., accuracy and precision for quantitative assays). As described in Section IV.A.4, these parameter estimates are used to make an objective decision to accept or reject the analytical method.

E. Stability. Drug stability in the study sample matrix is of paramount concern and must therefore be subject to investigation during Pre-Study Validation. Evaluation of the stability of the analyte should be performed in a representative matrix with assessment of freeze-thaw stability, short-term room temperature stability and long-term storage stability. Freeze-thaw stability should be evaluated at both -70°C and -20°C , if appropriate, since many macromolecules are unstable at -20°C . Without exception, however, the stability protocol should be based on assessments in which samples are subjected to conditions identical to those expected for the study sample conditions. Thus, whereas suggestions have been made that three freeze-thaw cycles should be routinely assessed, in certain applications a larger or smaller number of cycles might be realized during the routine use of the bioanalytical method. Therefore, stability testing should not be fixed at three cycles, but rather the maximum number of freeze-thaw cycles that samples will experience during the routine use of the method.

An additional concern with immunoassays is reagent stability. For example, the enzyme that is conjugated to the antibody in a typical ELISA will inevitably lose its enzymatic activity over time. Therefore during Pre-Study Validation, documentation should be made of the conditions under which the principle reagents maintain sufficient stability to meet the basic requirements of assay performance.

3. In-Study Validation

In-Study Validation should focus on the routine use of the calibration model and QC samples as described in Section IV.B. A calibration curve and a set of QC samples are recommended for each bioanalytical batch. Most laboratories consider a 96-well plate to constitute a bioanalytical batch for an ELISA, but several plates may be grouped into a single batch as long as the larger batch size is covered during Pre-Study Validation.

B. Anti-Drug Antibody Assays

The assays used to monitor subjects for the presence of anti-drug antibodies represent a unique class of bioanalytical procedures that have distinct requirements for assay validation. While there are several platforms currently used for the detection of anti-drug antibodies, these platforms generally fall into one of two classifications: immunoassays or bioassays. The major distinction between these classifications is that immunoassays serve to measure antibodies that bind to a drug (i.e., "binding" antibodies), while bioassays are designed to measure antibodies that neutralize the biological effect of

²⁷ For guidance in evaluating the aptness of calibration models, the reader is referred to ref 5.

a drug (i.e., “neutralizing” antibodies). The foregoing discussion focuses on immunoassay measurements of binding antibodies, whereas methods that address neutralizing antibodies are discussed in the Bioassay section of this report.

1. Bioanalytical Method Establishment

In addition to procuring an appropriate reference standard and documenting its identity and purity (see section on Standard Immunoassays), Bioanalytical Method Establishment should focus on procuring a positive control antibody. The positive control antibody should ideally emulate the antibodies present in the test samples as closely as possible, and for assays intended to support studies in man, an antibody-positive clinical sample would be optimal. However, early in the development of a drug, it is rare that such a reagent exists, and even in mature clinical programs, sufficient quantities of human anti-sera can seldom be procured. Some laboratories therefore depend on the use of pooled antibody-positive samples from non-human primates that were exposed to the drug during the preclinical stages of development. The next appropriate choice would be an affinity purified polyclonal antibody from a more evolutionarily distant mammal (such as a mouse or rabbit), followed by a monoclonal antibody that is known to bind the drug molecule.

2. Pre-Study Validation

A. Specificity. In anti-drug antibody assays, the analyte corresponds to a polygonal mixture of immunoglobulins having differential affinities for a multitude of binding sites (epitopes) on the drug. Given that the analyte is an extraordinarily heterogeneous entity, it is a forgone conclusion that the bioanalytical method will have a poor capacity to differentiate and quantitate the analyte in the presence of other constituents in a sample. For this reason, evaluations of method specificity defer to basic assessments of the certainty that an observed positive response corresponds to drug-binding antibodies.

The bioanalytical result in an immunoassay for anti-drug antibodies generally corresponds to a simple direct comparison of an instrument response to a test sample versus a set of reference samples. For example, where the ELISA format is used to measure anti-drug antibodies, the bioanalytical result is the units of optical density (OD) for the test sample relative to the OD value with a negative control reference sample where surface plasmon resonance is used to measure antibodies (X), the bioanalytical result is a change in the optical signal. In interpreting whether a sample is positive for the presence of anti-drug antibodies, analysts typically apply the following criteria:

1. The observed instrument response must be greater than the cutpoint.²⁸
2. The observed instrument response must be abrogated by addition of soluble drug to the assay.

²⁸ The cutpoint generally corresponds to measurements of samples from an untreated population, typically with a sample size of at least 30 individual subjects. The cutpoint is commonly defined as the average plus 3 standard deviations. (For a sample size of 30 subjects, 3 standard deviations corresponds to a 0.5% probability of reaching a false positive result.)

3. The onset of the antibody positive result should be correlated with exposure to the drug. (A typical criterion is that the instrument response ratio of post-dose:pre-dose (baseline) samples is ≥ 2.0).

4. Data from experimental controls should prove that relevant antibodies produce a positive result in the assay and irrelevant antibodies produce a negative result.

B. Assay Calibration. Calibration of anti-drug antibody assays is an area of considerable debate (7). Thus, while it is broadly considered to be impossible to procure a reference sample that universally matches study samples in the range of affinities and classes of immunoglobulins, regulators have expressed frustration with bioanalytical data that are expressed in arbitrary units such as titers. The principle frustrations with arbitrary units are two-fold: 1) it is not possible to evaluate whether a particular method has acceptable sensitivity; and 2) antibody data for a particular pharmaceutical product cannot be compared to data from other products. Given these frustrations, it is possible that drug companies will soon be required to make a good faith effort to develop antibody reference standards (i.e., anti-drug immunoglobulins purified from anti-serum) for calibrating anti-drug antibody assays. In this case, data generated in anti-drug antibody techniques may eventually be routinely reported in mass unit concentrations (e.g., ng/mL of IgG).

At present, the vast majority of anti-drug antibody bioanalytical methods generate semi-quantitative data in which antibody concentrations are reported in relative or arbitrary units. Thus, data are typically reported as titers, the reciprocal of the highest dilution of a sample in which the instrument response is greater than the cutpoint response. Where a positive control antibody is available, a pseudo-calibration curve can be generated by assay of serially diluted aliquots, and the concentration-response curve can be used to ensure that the method performs in a reproducible manner.

C. Precision, Accuracy and Recovery. Pre-Study Validation tests for the precision of bioanalytical methods for anti-drug antibody measurement typically focus on whether positive and negative control samples are measured with suitable reproducibility. Given that most anti-drug antibody assays generate semi-quantitative data, accuracy and recovery are not typically measured.

D. Quality Control (QC) Samples. The types of QC samples that are typically used in quantitative bioanalytical methods are seldom used in anti-drug antibody assays. Rather, determination of whether a particular batch has met performance acceptance criteria is commonly based on assay results with positive and negative control samples. As mentioned previously, the positive control generally corresponds to an antibody that is known to bind to the drug molecule. The negative control is typically a panel of serum samples from untreated subjects. An additional negative control that should be considered is assay of the test samples in the absence of the antigen, which serves to measure nonspecific binding of immunoglobulins in the test samples. Although the guidelines for Pre-Study Validation described in Section IV.A are not directly applicable to anti-drug antibody assays (since the guidelines specifically address quantitative methods), some of the guiding principles should be applied during the Pre-Study Validation of anti-drug antibody assays. For example, Target Acceptance Limits should be defined for the performance of the control samples, an Experimental Design

should be implemented that outlines the conduct of the Pre-Study Validation experiments, and statistical criteria should be established *a priori* for concluding that the method is either acceptable or unacceptable.

E. Stability. The stability tests described for Standard Immunoassays are applicable to anti-drug antibody bioanalytical methods.

3. In-Study Validation

Existing guidance documents for assay validation emphasize assessment of the calibration model during In-Study Validation. However, most anti-drug antibody assays generate semi-quantitative results, and the calibration model is therefore of lesser concern during In-Study Validation. Rather, In-Study Validation should focus on the performance of positive and negative controls. The QC samples described in Section III.B.2.D are well suited to serve in this capacity, the objective being to use the data generated with these controls to determine whether the results from a particular batch meet pre-defined acceptance criteria.

C. Bioassays (Cell-Based Assays/Activity-Based Assays)

A bioassay is a bioanalytical technique that uses a living system to measure the biological activity of a drug. In general, the biological activity relates to the intended therapeutic use of the drug, but in some cases may address drug toxicity or side effects. The primary focus of this section is validation of bioassays involving *in vitro* systems, for example, cell-based or activity-based assays. Although it is acknowledged that *in vivo* bioassays play an important role in pharmaceutical analytics, methodologies involving animals are considered beyond the scope of this document.

For the purpose of validation, an *in vitro* bioassay generally involves the use of an established (immortal) cell line to measure a discrete response such as cellular proliferation, survival, or differentiation. Such cell-based bioassays may be used for quantitative and semi-quantitative purposes. Among the most frequently used quantitative applications are estimations of drug potency to support manufacturing and measurements of concentrations of biologically active drug in *ex vivo* samples to support preclinical and clinical programs. Semi-quantitative applications include evaluation of drug neutralization in *ex vivo* samples and high-throughput screening to support drug discovery.

1. Bioanalytical Method Establishment

In addition to procuring an appropriate reference standard and documenting its identity and purity (see section on Standard Immunoassays), Bioanalytical Method Establishment should focus on documenting characteristics of the cell line. Examples of important characteristics may include, but are not limited to: origin of the cell line, culture history, subcloning history, morphology, surface markers, and the presence of microbial contamination. In the same context, it is advisable to research the effect of age of the cell culture on the response measured in the assay.

2. Pre-Study Validation

A. Specificity. Most cell lines proliferate, differentiate and/or senesce in response to a plethora of biomolecules.

Macromolecular drugs often undergo some form of metabolism *in vivo*, resulting in the generation of metabolites that may possess some biological activity. Moreover, biopharmaceutical therapies occasionally result in up- or down-regulation of biomolecules that elicit positive responses in the bioassay. Therefore it deserves to be emphasized that bioanalytical methods based on bioassays are rarely specific for the analyte of interest. In recognition of this limitation, during Pre-Study Validation, it is recommended to evaluate the extent of interference that occurs due to nonspecificity. Since a host of tissue factors may affect the cells, and endogenous concentrations of these factors may vary among individual subjects, baseline samples from a number of representative subjects should be studied for nonspecific induction of the cellular response. When bioassays are used to support clinical studies, the baseline samples will optimally be obtained from the target patient population.

B. Assay Calibration. For bioassays that derive quantitative measures of analyte concentration, the description of calibration methods provided in the section on Standard Immunoassays is applicable. In cases where the bioassay is applied to measurement of drug neutralization, assay calibration typically focuses on samples that are amended with the drug, commonly at the EC_{50} . When drug-neutralizing substances are present in the serum, the cellular response to the drug is selectively diminished. Therefore, the endpoint in the neutralizing assay is typically the ratio of the cellular response at the EC_{50} in the presence and absence of neutralizing substances. The robustness of this response must be thoroughly investigated during Pre-Study Validation.

C. Precision, Accuracy, and Recovery. For bioassays that derive quantitative measures of analyte concentration, the description of precision, accuracy and recovery methods provided in the section on Standard Immunoassays is applicable. For semi-quantitative bioassays, precision is of principle concern, with precision testing typically focusing on the performance of positive and negative control samples.

D. Quality Control (QC) Samples. For bioassays that derive quantitative measures of analyte concentration, the description of QC sample methodology provided in the section on Standard Immunoassays is applicable. For semi-quantitative bioassays, determination of whether a particular batch has met performance acceptance criteria is commonly based on assay results with positive and negative control samples, the objective being to use the data generated with these controls to determine whether the results from a particular batch meet pre-defined acceptance criteria (see Section III.B.2.D).

E. Stability. The stability tests described for Standard Immunoassays are applicable to bioassays.

3. In-Study Validation

For bioassays that derive quantitative measures of analyte concentration, the description of In-Study Validation provided in the section on Standard Immunoassays is applicable. For bioassays that generate semi-quantitative results, In-Study Validation should focus on the use of positive and negative controls. As described for anti-drug antibody assays (Section III.B.3), these controls are used during In-Study Validation to assess the reliability of results from individual batches.

D. Biomarkers

A biomarker is a measurable biological response that is related to the progression of a disease (8). When a biomarker is validated and linked to a clinical outcome according to appropriate criteria, it can be used to determine the drug efficacy or toxicity. While biomarkers can be small molecules (e.g., steroids and ion cofactors) or large molecules (nucleic acids, receptors, enzymes, proteins, polysaccharides, phospholipids, etc.), the foregoing discussion focuses on macromolecular forms, which are chiefly analyzed by immunochemical, cell-based, or enzyme assays.

1. Bioanalytical Method Establishment

Where the intended reference standard is a recombinantly expressed protein, it should be noted that such proteins often have glycosylation patterns that are distinct from their endogenous counterparts, and an investigation should therefore be undertaken to determine whether the reference material produces a different assay response than the endogenous biomarker. If assay responses are known or suspected to be different for reference and endogenous materials, it should be acknowledged that the assay results support the determination of trends, but not absolute concentrations of the biomarker.

In the ideal case, purified endogenous protein from the host species is used as an authentic reference standard. In such a case, the reference standard should be thoroughly characterized in terms of analytical purity. When commercial sources are used, detailed information regarding product characterization should be pursued with the vendor.

It should be established whether data are available on the properties of the reference standard relative to a primary WHO standard (or other standardized programs), information that will aid in determining whether it is appropriate to compare results across different studies or laboratories. Finally, it should be acknowledged that some biomarkers cannot be well defined because of structural uncertainty and/or lack of global standardization (9,10); for these biomarkers, quantitation methods will be relative rather than absolute.

2. Pre-Study Validation

A. Specificity. Verification that a biomarker assay is specific for the intended analyte poses a formidable challenge. The vast majority of biomarkers are endogenous molecules present at some baseline concentration in the biological matrix of interest (e.g., plasma, serum, urine, or cerebral spinal fluid). Baseline concentrations typically vary among individuals, with the greatest extent of variation occurring in diseased subjects. For macromolecular biomarkers, it is often difficult, if not impossible to distinguish low levels of the endogenous molecule from nonspecific measurements in the assay due to matrix interference. Given these significant caveats, biomarker assays are typically suitable for measuring relative, rather than absolute concentrations, the appropriate endpoint being assessment of post-treatment increases (or decreases) relative to pre-treatment values.

B. Assay Calibration. Whether aimed at determining absolute or relative concentrations, it is helpful during the development of a biomarker assay to incorporate screening of samples from a large number of healthy subjects as well as

subjects diagnosed with the disease. These screening results indicate the population distribution of biomarker concentrations, information that is pertinent to the selection of the control biological matrix used to prepare QC samples. In addition, the screening results can be used to optimize the design of the calibration curve. When endogenous levels are low in healthy controls and increase with disease, the calibration curve will ideally cover the lower normal levels, since this will facilitate quantitative assessment of the transitions between normal and disease state.

Preparation of the calibration standards in the biological matrix of interest has the inherent advantage of controlling for nonspecific signal with the bioanalytical method due to matrix interference. However, unless the biomarker shows a very large increase in expression over baseline concentrations, baseline concentrations may obscure the bioanalytical response to the up-regulated biomarker, resulting in an insensitive method. Moreover, such an approach is unsuitable for methods aimed at measuring biomarkers subject to down-regulation. Some laboratories purport affinity-stripped matrices to be the ideal medium for preparing calibration standards, since these have the added advantage of reduced interference due to the baseline concentrations of the endogenous biomarker. However, given that interference in a biomarker assay is seldom completely eliminated (neither nonspecific nor specific), even this approach may not yield a method with adequate sensitivity. Therefore, in most cases, biomarker bioanalytical methods incorporate the use of alternative matrices for calibration standard preparation, such as a buffer amended with serum proteins or a surrogate matrix from another animal species.

C. Precision, Accuracy, and Recovery. For biomarker assays that derive quantitative measures of analyte concentration, the description of precision, accuracy and recovery methods provided in the section on Standard Immunoassays is applicable. For semi-quantitative bioassays, precision is of principle concern, with precision testing typically focused on the performance of positive control samples.

D. Quality Control (QC) Samples. There are two guiding principles in the establishment of QC samples for a biomarker assay: 1) they should be prepared in a medium that mimics the matrix of the study samples; and 2) the concentrations should reflect those expected in the study samples. To fulfill these criteria, a pooled lot of matrix should be ideally used, and a set of QC samples should be prepared by spiking the analyte into the matrix at discrete concentrations. In cases where biomarker levels are reduced as a result of treatment with a biopharmaceutical, it may be necessary to prepare the QC samples in an alternate matrix as described above.

E. Stability. The stability tests described for Standard Immunoassays are applicable to biomarker assays.

3. In-Study Validation

For biomarker assays that derive quantitative measures of analyte concentration, the description of In-Study Validation provided in the section on Standard Immunoassays is applicable. For biomarker assays that generate semi-quantitative results, In-Study Validation should focus on the use of positive and negative controls as described in Section III.B.3.

IV. VALIDATION AND ACCEPTANCE CRITERIA

As evidenced from these workshop proceedings, there are several types of bioanalytical methods that are being routinely applied by pharmaceutical scientists to analyze macromolecules. Some of these methods derive semi-quantitative endpoints (e.g., titers of anti-drug antibodies), whereas others seek to obtain quantitative determinations (e.g., drug concentrations in pharmacokinetic samples). The purpose of this section is to address the latter.

A. Pre-Study Validation Criteria

The goal of Pre-Study Validation is to document that a bioanalytical method routinely produces results that are reliable. Accordingly, key considerations are that the method delivers suitable accuracy, repeatability (i.e., within-batch precision), and intermediate precision (i.e., between-batch precision).

To allow an objective evaluation of method acceptability during Pre-Study Validation, bioanalytical specifications must first be defined. Thus, it is recommended that prior to embarking on the Pre-Study Validation experiments, specification be made of the components listed below.

1. Target Acceptance Limits

The Target Acceptance Limits define the minimal performance required of a validated method. These limits should be established in collaboration with other scientists (e.g., pharmacokineticists) and in consideration of guidance documents, scientific publications and previous laboratory experience. The consensus opinion at the workshop was that the minimal acceptance limits for ligand-binding assays should be set at 30% for accuracy (mean bias) and precision, with limits greater than 30% being permissible if agreed upon by the end-users of the analytical data.

2. Experimental Design

The Experimental Design outlines the conduct of the Pre-Study Validation experiments. This represents a blueprint for evaluating the key method performance factors described in previous sections of this report, including specificity, the calibration model, precision, accuracy, recovery, dilutional linearity, and stability. Moreover, the Experimental Design should describe the calibrator concentrations, validation QC sample concentrations, and the number of validation batches and replicates. Regarding the latter issue, the following recommendations are made:

1. Since ligand-binding assays often have relatively high inter-assay imprecision, it is recommended that at least 6 validation batches be performed over several days. For bioanalytical methods with large bias or intermediate precision (e.g., absolute error $\geq 80\%$ of the Target Acceptance Limit), the number of batches should be increased to at least 10.

2. At least 3 sets of validation QC samples should be included in each batch, with each set of samples analyzed in at least duplicate. Each set should contain samples that span the quantitative range of the method (e.g., LLOQ, midrange, and ULOQ).

3. Procedures for Statistical Analysis

The Procedures for Statistical Analysis refer to the statistical methods used to analyze data from the Pre-Study Validation experiments. As described in detail by Findlay *et al.* (3), for validation experiments in which replicate measurements are made over multiple bioanalytical batches, accuracy and precision of a method are commonly estimated using a one-way random effects model. Although statistical software currently exists for performing these calculations, it is recognized that there continues to be a need for software designed specifically for analyzing data from validation experiments.

4. Statistical Acceptance Criteria

The Statistical Acceptance Criteria used during Pre-Study Validation constitute the basis for concluding that a bioanalytical method is either acceptable or unacceptable for its intended use. For the purpose of objectively accepting or rejecting a particular method, Statistical Acceptance Criteria are commonly applied to assay parameters such as accuracy and precision. These criteria fall into two categories:

1. Criteria for comparing estimates of mean assay parameter values with the Target Acceptance Limits.
2. Criteria for comparing statistical confidence intervals for the true assay parameter values with the Target Acceptance Limits.

Standard statistical tests designed to reject a null hypothesis are inappropriate for validation experiments because the desired outcome is to accept a method. Consequentially, a common practice is to default to category 1 above, namely to use Statistical Acceptance Criteria that directly compare estimates of mean assay parameters from Pre-Study Validation experiments with the Target Acceptance Limits (e.g., accuracy and precision each average $\leq 30\%$). The difficulty with this approach is that it does not limit the chances of falsely accepting an unsuitable method or falsely rejecting a suitable one. Because validation data are often limited (N is relatively small), the category 1 approach is often associated with a relatively high risk of falsely concluding that a bioanalytical method will achieve the desired performance in routine use. For these reasons, confidence intervals and equivalence testing procedures (category 2 above) have been proposed (11–12).

A shortcoming of both category 1 and 2 criteria is that the underlying basis for deeming a method to be suitable or unsuitable during Pre-Study Validation is inconsistent with the In-Study Validation recommendation to use the “4-6-30 rules” (see below). An alternative approach that provides greater consistency between Pre-Study and In-Study Validation results is to define Pre-Study Statistical Acceptance Criteria as limits on the prediction interval for total measurement error (13,14). This approach combines both accuracy and precision into a single integrated assessment of method suitability in a manner similar to the ‘4-6-30 rule’. Because of concerns about the increased computational requirements of the prediction interval approach, a simple compromise is proposed that achieves some of the goals of the total error approach and maintains consistency with the In-Study acceptance criteria. In this approach, the Pre-Study Validation criteria for method acceptance requires the sum of the absolute

mean bias and the intermediate precision to be $\leq 30\%$ (see Table II).

B. In-Study Validation Criteria

The following recommendations apply to acceptance of an In-Study validation assay run:

1. Adopt a "4-6-30 rule" as the criteria for accepting In-Study runs (4 out of 6 QC results must be within 30% of their respective nominal values).
2. For each QC concentration level, at least 50% of the results must be within 30% of the respective nominal value.
3. Since QC results must be in-range at all concentration levels, results cannot be reported from a truncated standard (calibration) curve.

Another point to consider is that standard curve parameters (e.g. maximum binding, nonspecific binding, ED_{20} , ED_{50} , ED_{80} , min-max signal, and signal-to-noise ratio) may be monitored to assess standard curve reproducibility. Moreover, although anchor points can be useful in calibration, it is inappropriate to report results beyond the validated range (i.e., outside the limits of quantitation).

V. CLOSING COMMENTS

This report has summarized the major issues discussed at the March 2000 workshop on "Bioanalytical Methods Validation for Macromolecules." Its purpose was to provide an initial set of guiding principles for the validation of bioanalytical methods used to support the preclinical and clinical stages of macromolecular drug development. Despite the substantial groundwork laid at the meeting, due to the diversity in topics, there was not adequate time for an in-depth discussion of all the analytical issues. It is clear that numerous unresolved issues concerning the validation and application of analytical methods for macromolecules warrant further discussion. These issues include, but are not limited to, questions such as: What is the optimal method to evaluate "parallelism?" What are the best approaches for nonlinear calibration? Are anchor calibrators necessary? What defines the quantitative range? Are specific recommendations needed for editing standard curves? What is the best approach to assess accuracy (mean bias) and imprecision? And what are the ap-

proaches for defining specificity (selectivity)? In addition to noting these unresolved issues, it should be emphasized again that the recommendations in this document do not necessarily reflect a concordance of viewpoints. It is the authors' hope that this report will inspire further discussion and serve as a foundation and framework for ongoing efforts to address the unresolved issues, as well as the points of controversy. Accordingly, several of this report's authors have joined with fellow bioanalytical scientists to establish the Ligand Binding Assay Bioanalytical Focus Group within the AAPS to provide a forum for subsequent discussions pertaining to the development of globally accepted best practices for the bioanalysis of macromolecules. For information about this group, the reader is advised to contact AAPS directly²⁹ or to read the posting on the Internet at www.aapspharmaceutica.com.

VI. GLOSSARY OF TERMS (DEFINITIONS)

4-6-20 QC rule: A batch (run) acceptance criterion widely used in the pharmaceutical industry, which requires that 4 out of 6 QC results be within $\pm 20\%$ of their respective nominal value. Recently, this rule was modified for small molecule chromatographic-based assays to require 67% (4 out of 6) of QC results to be within 20% of their respective nominal values; 33% of the results (not all replicates at the same concentration) may be outside the $\pm 20\%$ of the nominal value. In this document, this rule has been modified for the bioanalysis of macromolecules to require 4 out of 6 of QC results to be within $\pm 30\%$ of their respective nominal value with at least 50% of the QC results in-range at each concentration level.

Anti-drug antibody: An antibody that binds to a drug.

Accuracy (ICH): The closeness of agreement between the value that is accepted either as a conventional true value or an accepted reference value and the value found. This is sometimes termed trueness.

Batch: Synonymous with run. A set of standard curve calibrators, validation samples, and/or quality control samples, and/or study samples that is analyzed in a single group.

Bias: Systematic difference between measured a test result and the theoretical true value (nominal). Bias is expressed either as a relative error (% RE) or as a ratio (% recovery).

Binding antibody: An antibody that binds to a drug but does not necessarily serve to neutralize the biological activity of the drug.

Bioassay: An analytical format in which the response variable is biological in nature (e.g., cell proliferation).

Biomarker: A characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacological responses to a therapeutic intervention.

Calibration curve: A functional relationship between the analyte concentration in the standards (calibrators) and the measured response. The calibration curve is used to estimate the analyte concentration in test samples by dose interpolation.

Table II. Total Error Values Applicable to Bioanalytical Method Validation^a

Accuracy (mean bias)	Intermediate precision	Total error
0	30	30
5	25	30
10	20	30
15	15	30
20	10	30
25	5	30
30	0	30

^aTotal error in this table is defined as the sum of the absolute mean bias and intermediate precision. This table provides a simple alternative to the more complex computation of a prediction interval for total error (3). This computation is being recommended due to the lack of universally available software.

²⁹ AAPS, 2107 Wilson Blvd, Suite 700, Arlington, VA 22201-3046.

Calibration standards: Samples having a known concentration of analyte that are used in an assay to gauge the relationship between system responses (e.g., absorption units) concentrations of an analyte.

Calibrator: A solution or biological matrix spiked with the analyte of interest with which test samples are compared for estimating the concentration of analyte. For the purposes of this document, calibrator and standard are synonymous terms.

Competitive assay: A type of ligand-binding assay format in which the analyte competes with a labeled form of the analyte (e.g., radiolabel, enzyme, and fluorescent label) for a fixed and limiting concentration of the binder, usually an antibody or antiserum.

Cross-validation: Validation that supports the use of two or more bioanalytical methods within the same study.

Cutpoint: The cutpoint generally corresponds to measurements of samples from an untreated population, typically with a sample size of at least 30 individual subjects. The cutpoint is commonly defined as the average plus 3 standard deviations.

Dilutional linearity: A condition in which dilution of a spiked sample does not result in biased measurement of the analyte concentration. Thus, when a spiked sample is serially diluted to result in a set of samples having analyte concentrations that fall within the quantitative range of the assay, the entire set of dilutions is measured with acceptable accuracy.

Full validation: A validation that includes evaluation of accuracy, precision, curve-fitting (model assessment), sensitivity, specificity, stability, etc.

Immunoassay: A type of ligand-binding assay in which an antibody or antiserum is used to as the specific binder.

Intermediate precision (ICH): Precision of repeated measurements within-laboratories taking into account all relevant sources of variation affecting the results (e.g., day, analyst, batch). Also referred to as inter-batch, inter-assay, and inter-run precision.

Ligand-binding assay: A type of assay format that depends on the specific binding of an analyte to another molecule, usually a macromolecule (biopolymer). This format typically involves reversible noncovalent interactions governed by the laws of mass action.

Limit of detection: The lowest concentration of analyte for which the response can be reliably distinguished from background noise.

Linearity (ICH): A condition in which test results are directly proportional to the concentration (amount) of analyte in the sample.

Lower limit of quantitation (ICH): The lowest concentration (amount) of analyte in a test sample that can be determined quantitatively with suitable accuracy (mean bias) and precision.

Macromolecule: A molecule having a mass greater than 1,000 daltons. Macromolecules are commonly biopolymers that have the potential to provoke an immune response. Due to their inherent molecular complexity, macromolecules are generally more difficult to characterize than conventional small molecule xenobiotics.

Matrix: The material, usually of biological origin, in which the analyte is contained.

Matrix Effect: Interference in an assay that is caused by adding the sample matrix. Commonly refers to analytical in-

terference produced by factors other than those that have physicochemical similarity to the analyte.

Neutralizing antibody: An antibody that binds to a drug in such a way as to inhibit its biological activity.

Noncompetitive assay: A type of ligand-binding assay format in which the analyte is detected using multiple binders (e.g., sandwich ELISA). In this binding format at least one of the binders is present in excess amount and one is labeled (e.g., enzyme, radiolabel, fluorescence) or modified in a specific manner (e.g., biotinylated) to permit detection of the binding reaction.

Nonspecific nonspecificity: Analytical interference caused by factors other than those that are related physicochemically to the analyte of interest, but which nevertheless affect the *in vitro* binding reaction. This type of nonspecificity is commonly referred to as matrix effects.

Parallelism: A condition in which dilution of test samples does not result in biased measurements of the analyte concentration. Thus, when a test sample is serially diluted to result in a set of samples having analyte concentrations that fall within the quantitative range of the assay, there is no apparent trend toward increasing or decreasing estimates of analyte concentration over the range of dilutions.

Partial validation: Modification of a full validation. Partial validations can range from analysis of a single batch to nearly full validation and are typically performed to support bioanalytical method transfer, platform changes, changes in assay range, changes in the matrix species of origin, selectivity in the presence of co-administered drugs, etc.

Precision (ICH): The closeness of agreement (degree of scatter) between a series of measurements obtained from multiple sampling of the same homogeneous sample under the prescribed conditions. Precision may be considered at three levels: repeatability, intermediate precision and reproducibility.

Pre-Study Validation: Procedures used before the analysis of study samples to establish that an bioanalytical method is suitable for its intended application.

Quality control (QC) samples: Pre-Study Validation and In-Study samples having a known concentration (nominal) of analyte that are treated as unknowns in an assay. During Pre-Study Validation, QC samples are used to generate information to demonstrate the method is suitable for its intended purpose. During In-Study runs, QC values are used as the basis for accepting and rejecting bioanalytical method batches.

Range (ICH): The interval between the upper and lower concentrations (amounts) of analyte in the sample for which it has been demonstrated that the analytical procedure has a suitable level of accuracy (mean bias), precision, and linearity.

Recovery (spike recovery): A measurement of the closeness of an observed result to its theoretical true value. Recovery is generally expressed as the percentage of the observed to the nominal (theoretical) concentration. Spike recovery relates to cases where the theoretical concentration corresponds to the concentration of analyte added to a sample by the analyst.

Repeatability (ICH): The precision under the same operating conditions over a short interval of time. Also termed intra-batch or intra-run precision.

Reproducibility (ICH): Precision of repeated measurements between laboratories. Also termed inter-laboratory precision. Usually applies to collaborative studies that involve the standardization of a bioanalytical method across multiple laboratories.

Response error relationship: The relationship between the variability in replicate response measurements (e.g., cpm, absorbance) and the mean response.

Robustness (ICH): A measure of a bioanalytical method's capacity to remain unaffected by small, but deliberate, variations in method parameters. Robustness provides an indication of a bioanalytical method's reliability during normal usage.

Run: Synonymous with batch. The collection of analytical samples covered by a single calibration curve and set of QC samples. The size of a bioanalytical run must be defined during Pre-Study Validation.

Selectivity: The extent to which a bioanalytical method can measure particular analyte(s) in a complex mixture without interference from other components of the mixture.

Specific nonspecificity: Analytical interference that is caused by substances in the test sample that have physicochemical similarity to the analyte of interest. Examples of such substances include metabolites, degraded forms of the analyte, isoforms, precursors, and structural variants that differ with regard to post-translational modification.

Specificity: The ability to unequivocally measure the analyte in the presence of other components that may be expected to be present in the biological specimen, including impurities, metabolites, and endogenous matrix components.

Total error: A term that describes the agreement between a measured test result and the theoretical true value. The term total error describes a combination of systematic (mean bias) and random (imprecision) error components. In some publications, the term total error is also defined to as accuracy (ISO).

Upper limit of quantitation: The highest concentration (amount) of analyte in a test sample that can be quantitatively determined with suitable accuracy (mean bias) and precision.

REFERENCES

1. V. P. Shah, K. K. Midha, S. Dighe, I. J. McGilveray, J. P. Skelley, A. Yacobi, T. Layloff, C. T. Viswanathan, C. E. Cook, R. D. Mc-

- Dowall, K. A. Pittman, and S. Spector. Analytical methods validation: bioavailability, bioequivalence and pharmacokinetic studies. *Pharm. Res.* **9**:588–592 (1992).
2. V. P. Shah, K. K. Midha, J. W. A. Findlay, H. M. Hill, J. D. Hulse, J. McGilveray, G. McKay, K. J. Miller, R. N. Patnaik, M. L. Powell, A. Tonelli, C. T. Viswanathan, and A. Yacobi. Bioanalytical methods validation—a revisit with a decade of progress. *Pharm. Res.* **17**:1551–1557 (2000).
3. J. W. Findlay, W. C. Smith, J. W. Lee, G. D. Nordblom, I. Das, B. S. DeSilva, M. N. Khan, and R. R. Bowsher. Validation of immunoassays for bioanalysis: a pharmaceutical industry perspective. *J. Pharm. Biomed. Anal.* **21**:1249–1273 (2000).
4. H. T. Karnes, G. Shiu, and V. P. Shah. Validation of bioanalytical methods. *Pharm. Res.* **8**:421–426 (1991).
5. J. Neter, W. Wasserman, and M. H. Kutner. Aptness of model and remedial measures. In *Applied Linear Regression Models*. Richard D. Irwin, Inc., Homewood, Illinois, 1983. pp. 109–146.
6. M. A. Takacs, S. J. Jacobs, R. M. Bordens, and S. J. Swanson. Detection and characterization of antibodies to PEG-IFN- α 2b using surface plasmon resonance. *J. Interferon Cytokine Res.* **19**:781–789 (1999).
7. [FDA] Food and Drug Administration. Center for Biologics Evaluation and Research, Meeting of the Biological Response Modifiers Advisory Committee. Thursday, July 15, 1999.
8. G. J. Downing. *Biomarkers and surrogate endpoints: clinical research and applications*. Elsevier Science, Amsterdam, The Netherlands 2000.
9. M. Wadhwa and R. Thorpe. Standardization and calibration of cytokine immunoassays: meeting report and recommendations. *Cytokine* **9**:791–793 (1997).
10. A. Ledur, C. Fitting, B. David, C. Hamberger, and J. M. Cavailon. Variable estimates of cytokine levels produced by commercial ELISA kits: results using international cytokine standards. *J. Immunol. Methods* **186**:171–179 (1995).
11. R. O. Kringle. An assessment of the 4–6–20 rule for acceptance of analytical runs in bioavailability, bioequivalence, and pharmacokinetic studies. *Pharm. Res.* **11**:556–560 (1994).
12. C. Hartmann, D. L. Massart, and R. D. McDowall. An analysis of the Washington Conference Report on bioanalytical method validation. *J. Pharm. Biomed. Anal.* **12**:1337–13343 (1994).
13. R. O. Kringle and R. C. Khan-Malek. A statistical assessment of the recommendations from a conference on analytical methods validation in bioavailability, bioequivalence, and pharmacokinetic studies. *Proceedings of the Biopharmaceutical Section of the American Statistical Association*, Alexandria, VA, August 13–18, 1994. 510–514.
14. P. Hubert, P. Chiap, J. Crommen, B. Boulanger, E. Chapuzet, N. Mercier, S. Bervoas-Martin, P. Chevalier, D. Grandjean, P. Lagorce, M. Lallier, M. C. Laparra, M. Laurentie, and J. C. Nivet. The SFSTP guide on the validation of chromatographic methods for drug analysis: from the Washington Conference to the laboratory. *Anal. Chim. Acta* **391**:135–148 (1999).