Guidance for Industry

Bioanalytical Method Validation

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Veterinary Medicine (CVM)
May 2001
BP

Guidance for Industry

Bioanalytical Method Validation

Additional copies are available from:

Drug Information Branch (HFD-210)
Center for Drug Evaluation and Research (CDER)
5600 Fishers Lane, Rockville, MD 20857 (Tel) 301-827-4573
Internet at http://www.fda.gov/cder/guidance/index.htm

or

Communications Staff (HFV-12)
Center for Veterinary Medicine (CVM)
7500 Standish Place, Rockville, MD 20855 (Tel) 301–594-1755
Internet at http://www.fda.gov/cvm

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Veterinary Medicine (CVM)
May 2001
BP

Table of Contents

GUIDANCE FOR INDUSTRY[1]

Bioanalytical Method Validation

I. INTRODUCTION

This guidance provides assistance to sponsors of investigational new drug applications (INDs), new drug applications (NDAs), abbreviated new drug applications (ANDAs), and supplements in developing bioanalytical method validation information used in human clinical pharmacology, bioavailability (BA), and bioequivalence (BE) studies requiring pharmacokinetic (PK) evaluation. This guidance also applies to bioanalytical methods used for non-human pharmacology/toxicology studies and preclinical studies. For studies related to the veterinary drug approval process, this guidance applies only to blood and urine BA, BE, and PK studies.

The information in this guidance generally applies to bioanalytical procedures such as gas chromatography (GC), high-pressure liquid chromatography (LC), combined GC and LC mass spectrometric (MS) procedures such as LC-MS, LC-MS-MS, GC-MS, and GC-MS-MS performed for the quantitative determination of drugs and/or metabolites in biological matrices such as blood, serum, plasma, or urine. This guidance also applies to other bioanalytical methods, such as immunological and microbiological procedures, and to other biological matrices, such as tissue and skin samples.

This guidance provides general recommendations for bioanalytical method validation. The recommendations can be adjusted or modified depending on the specific type of analytical method used.

II. BACKGROUND

[1] This guidance has been prepared by the Biopharmaceutics Coordinating Committee in the Center for Drug Evaluation and Research (CDER) in cooperation with the Center for Veterinary Medicine (CVM) at the Food and Drug Administration.

This guidance has been developed based on the deliberations of two workshops: (1) Analytical Methods Validation: Bioavailability, Bioequivalence, and Pharmacokinetic Studies (held on December 3**B**5, 1990[2]) and (2) Bioanalytical Methods Validation **C** A Revisit With a Decade of Progress (held on January 12**B**14, 2000[3]).

Selective and sensitive analytical methods for the quantitative evaluation of drugs and their metabolites (analytes) are critical for the successful conduct of preclinical and/or biopharmaceutics and clinical pharmacology studies. Bioanalytical method validation includes all of the procedures that demonstrate that a particular method used for quantitative measurement of analytes in a given biological matrix, such as blood, plasma, serum, or urine, is reliable and reproducible for the intended use. The fundamental parameters for this validation include (1) accuracy, (2) precision, (3) selectivity, (4) sensitivity, (5) reproducibility, and (6) stability. Validation involves documenting, through the use of specific laboratory investigations, that the performance characteristics of the method are suitable and reliable for the intended analytical applications. The acceptability of analytical data corresponds directly to the criteria used to validate the method.

Published methods of analysis are often modified to suit the requirements of the laboratory performing the assay. These modifications should be validated to ensure suitable performance of the analytical method. When changes are made to a previously validated method, the analyst should exercise judgment as to how much additional validation is needed. During the course of a typical drug development program, a defined bioanalytical method undergoes many modifications. The evolutionary changes to support specific studies and different levels of validation demonstrate the validity of an assay's performance. Different types and levels of validation are defined and characterized as follows:

A. Full Validation

- Full validation is important when developing and implementing a bioanalytical method for the first time.
- Full validation is important for a new drug entity.
- A full validation of the revised assay is important if metabolites are added to an existing assay for quantification.

B. Partial Validation

Partial validations are modifications of already validated bioanalytical methods. Partial validation can range from as little as one intra-assay accuracy and precision determination to a nearly full

[2] Workshop Report: Shah, V.P. et al., *Pharmaceutical Research*: 1992; 9:588-592.
[3] Workshop Report: Shah, V.P. et al., *Pharmaceutical Research*: 2000; 17:in press.

validation. Typical bioanalytical method changes that fall into this category include, but are not limited to:

- Bioanalytical method transfers between laboratories or analysts
- Change in analytical methodology (e.g., change in detection systems)
- Change in anticoagulant in harvesting biological fluid
- Change in matrix within species (e.g., human plasma to human urine)
- Change in sample processing procedures
- Change in species within matrix (e.g., rat plasma to mouse plasma)
- Change in relevant concentration range
- Changes in instruments and/or software platforms
- Limited sample volume (e.g., pediatric study)
- Rare matrices
- Selectivity demonstration of an analyte in the presence of concomitant medications
- Selectivity demonstration of an analyte in the presence of specific metabolites

C. Cross-Validation

Cross-validation is a comparison of validation parameters when two or more bioanalytical methods are used to generate data within the same study or across different studies. An example of cross-validation would be a situation where an original validated bioanalytical method serves as the *reference* and the revised bioanalytical method is the *comparator*. The comparisons should be done both ways.

When sample analyses within a single study are conducted at more than one site or more than one laboratory, cross-validation with spiked matrix standards and subject samples should be conducted at each site or laboratory to establish interlaboratory reliability. Cross-validation should also be considered when data generated using different analytical techniques (e.g., LC-MS-MS vs. ELISA[4]) in different studies are included in a regulatory submission.

All modifications should be assessed to determine the recommended degree of validation. The analytical laboratory conducting pharmacology/toxicology and other preclinical studies for regulatory submissions should adhere to FDA's Good Laboratory Practices (GLPs)[5] (21 CFR part 58) and to sound principles of quality assurance throughout the testing process. The bioanalytical method for human BA, BE, PK, and drug interaction studies must meet the criteria in 21 CFR 320.29. The analytical laboratory should have a written set of standard operating procedures (SOPs) to ensure a complete system of quality control and assurance. The SOPs should cover all aspects of analysis from the time the sample is collected and reaches the laboratory until the results of the analysis are reported. The SOPs also should include record keeping, security and chain of sample custody

[4] Enzyme linked immune sorbent assay

[5] For the Center for Veterinary Medicine, all bioequivalence studies are subject to Good Laboratory Practices.

(accountability systems that ensure integrity of test articles), sample preparation, and analytical tools such as methods, reagents, equipment, instrumentation, and procedures for quality control and verification of results.

The process by which a specific bioanalytical method is developed, validated, and used in routine sample analysis can be divided into (1) reference standard preparation, (2) bioanalytical method development and establishment of assay procedure, and (3) application of validated bioanalytical method to routine drug analysis and acceptance criteria for the analytical run and/or batch. These three processes are described in the following sections of this guidance.

III. REFERENCE STANDARD

Analysis of drugs and their metabolites in a biological matrix is carried out using samples spiked with calibration (reference) standards and using quality control (QC) samples. The purity of the reference standard used to prepare spiked samples can affect study data. For this reason, an authenticated analytical reference standard of known identity and purity should be used to prepare solutions of known concentrations. If possible, the reference standard should be identical to the analyte. When this is not possible, an established chemical form (free base or acid, salt or ester) of known purity can be used. Three types of reference standards are usually used: (1) certified reference standards (e.g., USP compendial standards); (2) commercially supplied reference standards obtained from a reputable commercial source; and/or (3) other materials of documented purity custom-synthesized by an analytical laboratory or other noncommercial establishment. The source and lot number, expiration date, certificates of analyses when available, and/or internally or externally generated evidence of identity and purity should be furnished for each reference standard.

IV. METHOD DEVELOPMENT: CHEMICAL ASSAY

The method development and establishment phase defines the chemical assay. The fundamental parameters for a bioanalytical method validation are accuracy, precision, selectivity, sensitivity, reproducibility, and stability. Measurements for each analyte in the biological matrix should be validated. In addition, the stability of the analyte in spiked samples should be determined. Typical method development and establishment for a bioanalytical method include determination of (1) selectivity, (2) accuracy, precision, recovery, (3) calibration curve, and (4) stability of analyte in spiked samples.

A. Selectivity

Selectivity is the ability of an analytical method to differentiate and quantify the analyte in the presence of other components in the sample. For selectivity, analyses of blank samples of the appropriate biological matrix (plasma, urine, or other matrix) should be obtained from at least

six sources. Each blank sample should be tested for interference, and selectivity should be ensured at the lower limit of quantification (LLOQ).

Potential interfering substances in a biological matrix include endogenous matrix components, metabolites, decomposition products, and in the actual study, concomitant medication and other exogenous xenobiotics. If the method is intended to quantify more than one analyte, each analyte should be tested to ensure that there is no interference.

B. Accuracy, Precision, and Recovery

The ***accuracy*** of an analytical method describes the closeness of mean test results obtained by the method to the true value (concentration) of the analyte. Accuracy is determined by replicate analysis of samples containing known amounts of the analyte. Accuracy should be measured using a minimum of five determinations per concentration. A minimum of three concentrations in the range of expected concentrations is recommended. The mean value should be within 15% of the actual value except at LLOQ, where it should not deviate by more than 20%. The deviation of the mean from the true value serves as the measure of accuracy.

The ***precision*** of an analytical method describes the closeness of individual measures of an analyte when the procedure is applied repeatedly to multiple aliquots of a single homogeneous volume of biological matrix. Precision should be measured using a minimum of five determinations per concentration. A minimum of three concentrations in the range of expected concentrations is recommended. The precision determined at each concentration level should not exceed 15% of the coefficient of variation (CV) except for the LLOQ, where it should not exceed 20% of the CV. Precision is further subdivided into within-run, intra-batch precision or repeatability, which assesses precision during a single analytical run, and between-run, inter-batch precision or repeatability, which measures precision with time, and may involve different analysts, equipment, reagents, and laboratories.

The ***recovery*** of an analyte in an assay is the detector response obtained from an amount of the analyte added to and extracted from the biological matrix, compared to the detector response obtained for the true concentration of the pure authentic standard. Recovery pertains to the extraction efficiency of an analytical method within the limits of variability. Recovery of the analyte need not be 100%, but the extent of recovery of an analyte and of the internal standard should be consistent, precise, and reproducible. Recovery experiments should be performed by comparing the analytical results for extracted samples at three concentrations (low, medium, and high) with unextracted standards that represent 100% recovery.

C. Calibration/Standard Curve

A calibration (standard) curve is the relationship between instrument response and known concentrations of the analyte. A calibration curve should be generated for each analyte in the

sample. A sufficient number of standards should be used to adequately define the relationship between concentration and response. A calibration curve should be prepared in the same biological matrix as the samples in the intended study by spiking the matrix with known concentrations of the analyte. The number of standards used in constructing a calibration curve will be a function of the anticipated range of analytical values and the nature of the analyte/response relationship. Concentrations of standards should be chosen on the basis of the concentration range expected in a particular study. A calibration curve should consist of a blank sample (matrix sample processed without internal standard), a zero sample (matrix sample processed with internal standard), and six to eight non-zero samples covering the expected range, including LLOQ.

1. Lower Limit of Quantification (LLOQ)

The lowest standard on the calibration curve should be accepted as the limit of quantification if the following conditions are met:

- The analyte response at the LLOQ should be at least 5 times the response compared to blank response.
- Analyte peak (response) should be identifiable, discrete, and reproducible with a precision of 20% and accuracy of 80-120%.

2. Calibration Curve/Standard Curve/Concentration-Response

The simplest model that adequately describes the concentration-response relationship should be used. Selection of weighting and use of a complex regression equation should be justified. The following conditions should be met in developing a calibration curve:

- #20% deviation of the LLOQ from nominal concentration

- #15% deviation of standards other than LLOQ from nominal concentration

At least four out of six non-zero standards should meet the above criteria, including the LLOQ and the calibration standard at the highest concentration. Excluding the standards should not change the model used.

D. Stability

Drug stability in a biological fluid is a function of the storage conditions, the chemical properties of the drug, the matrix, and the container system. The stability of an analyte in a particular matrix and container system is relevant only to that matrix and container system and should not be extrapolated to other matrices and container systems. Stability procedures should evaluate the stability of the analytes during sample collection and handling, after long-term (frozen at the

intended storage temperature) and short-term (bench top, room temperature) storage, and after going through freeze and thaw cycles and the analytical process. Conditions used in stability experiments should reflect situations likely to be encountered during actual sample handling and analysis. The procedure should also include an evaluation of analyte stability in stock solution.

All stability determinations should use a set of samples prepared from a freshly made stock solution of the analyte in the appropriate analyte-free, interference-free biological matrix. Stock solutions of the analyte for stability evaluation should be prepared in an appropriate solvent at known concentrations.

1. Freeze and Thaw Stability

Analyte stability should be determined after three freeze and thaw cycles. At least three aliquots at each of the low and high concentrations should be stored at the intended storage temperature for 24 hours and thawed unassisted at room temperature. When completely thawed, the samples should be refrozen for 12 to 24 hours under the same conditions. The freeze–thaw cycle should be repeated two more times, then analyzed on the third cycle. If an analyte is unstable at the intended storage temperature, the stability sample should be frozen at -70^0C during the three freeze and thaw cycles.

2. Short-Term Temperature Stability

Three aliquots of each of the low and high concentrations should be thawed at room temperature and kept at this temperature from 4 to 24 hours (based on the expected duration that samples will be maintained at room temperature in the intended study) and analyzed.

3. Long-Term Stability

The storage time in a long-term stability evaluation should exceed the time between the date of first sample collection and the date of last sample analysis. Long-term stability should be determined by storing at least three aliquots of each of the low and high concentrations under the same conditions as the study samples. The volume of samples should be sufficient for analysis on three separate occasions. The concentrations of all the stability samples should be compared to the mean of back-calculated values for the standards at the appropriate concentrations from the first day of long-term stability testing.

4. Stock Solution Stability

The stability of stock solutions of drug and the internal standard should be evaluated at room temperature for at least 6 hours. If the stock solutions are refrigerated or frozen

for the relevant period, the stability should be documented. After completion of the desired storage time, the stability should be tested by comparing the instrument response with that of freshly prepared solutions.

5. Post-Preparative Stability

The stability of processed samples, including the resident time in the autosampler, should be determined. The stability of the drug and the internal standard should be assessed over the anticipated run time for the batch size in validation samples by determining concentrations on the basis of original calibration standards.

Although the traditional approach of comparing analytical results for stored samples with those for freshly prepared samples has been referred to in this guidance, other statistical approaches based on confidence limits for evaluation of an analyte=s stability in a biological matrix can be used. SOPs should clearly describe the statistical method and rules used. Additional validation may include investigation of samples from dosed subjects.

E. Principles of Bioanalytical Method Validation and Establishment

- The fundamental parameters to ensure the acceptability of the performance of a bioanalytical method validation are accuracy, precision, selectivity, sensitivity, reproducibility, and stability.

- A specific, detailed description of the bioanalytical method should be written. This can be in the form of a protocol, study plan, report, and/or SOP.

- Each step in the method should be investigated to determine the extent to which environmental, matrix, material, or procedural variables can affect the estimation of analyte in the matrix from the time of collection of the material up to and including the time of analysis.

- It may be important to consider the variability of the matrix due to the physiological nature of the sample. In the case of LC-MS-MS-based procedures, appropriate steps should be taken to ensure the lack of matrix effects throughout the application of the method, especially if the nature of the matrix changes from the matrix used during method validation.

- A bioanalytical method should be validated for the intended use or application. All experiments used to make claims or draw conclusions about the validity of the method should be presented in a report (method validation report).

- Whenever possible, the same biological matrix as the matrix in the intended samples should be used for validation purposes. (For tissues of limited availability, such as bone marrow, physiologically appropriate proxy matrices can be substituted.)

- The stability of the analyte (drug and/or metabolite) in the matrix during the collection process and the sample storage period should be assessed, preferably prior to sample analysis.

- For compounds with potentially labile metabolites, the stability of analyte in matrix from dosed subjects (or species) should be confirmed.

- The accuracy, precision, reproducibility, response function, and selectivity of the method for endogenous substances, metabolites, and known degradation products should be established for the biological matrix. For selectivity, there should be evidence that the substance being quantified is the intended analyte.

- The concentration range over which the analyte will be determined should be defined in the bioanalytical method, based on evaluation of actual standard samples over the range, including their statistical variation. This defines the *standard curve.*

- A sufficient number of standards should be used to adequately define the relationship between concentration and response. The relationship between response and concentration should be demonstrated to be continuous and reproducible. The number of standards used should be a function of the dynamic range and nature of the concentration-response relationship. In many cases, six to eight concentrations (excluding blank values) can define the standard curve. More standard concentrations may be recommended for nonlinear than for linear relationships.

- The ability to dilute samples originally above the upper limit of the standard curve should be demonstrated by accuracy and precision parameters in the validation.

- In consideration of high throughput analyses, including but not limited to multiplexing, multicolumn, and parallel systems, sufficient QC samples should be used to ensure control of the assay. The number of QC samples to ensure proper control of the assay should be determined based on the run size. The placement of QC samples should be judiciously considered in the run.

- For a bioanalytical method to be considered valid, specific acceptance criteria should be set in advance and achieved for accuracy and precision for the validation of QC samples over the range of the standards.

F. Specific Recommendations for Method Validation

- The matrix-based standard curve should consist of a minimum of six standard points, excluding blanks, using single or replicate samples. The standard curve should cover the entire range of expected concentrations.

- Standard curve fitting is determined by applying the simplest model that adequately describes the concentration-response relationship using appropriate weighting and statistical tests for *goodness of fit*.

- LLOQ is the lowest concentration of the standard curve that can be measured with acceptable accuracy and precision. The LLOQ should be established using at least five samples independent of standards and determining the coefficient of variation and/or appropriate confidence interval. The LLOQ should serve as the lowest concentration on the standard curve and should not be confused with the limit of detection and/or the low QC sample. The highest standard will define the upper limit of quantification (ULOQ) of an analytical method.

- For validation of the bioanalytical method, accuracy and precision should be determined using a minimum of five determinations per concentration level (excluding blank samples). The mean value should be within ±15% of the theoretical value, except at LLOQ, where it should not deviate by more than ±20%. The precision around the mean value should not exceed 15% of the CV, except for LLOQ, where it should not exceed 20% of the CV. Other methods of assessing accuracy and precision that meet these limits may be equally acceptable.

- The accuracy and precision with which known concentrations of analyte in biological matrix can be determined should be demonstrated. This can be accomplished by analysis of replicate sets of analyte samples of known concentrations **C** QC samples **C** from an equivalent biological matrix. At a minimum, three concentrations representing the entire range of the standard curve should be studied: one within 3x the lower limit of quantification (LLOQ) (low QC sample), one near the center (middle QC), and one near the upper boundary of the standard curve (high QC).

- Reported method validation data and the determination of accuracy and precision should include all outliers; however, calculations of accuracy and precision excluding values that are statistically determined as outliers can also be reported.

- The stability of the analyte in biological matrix at intended storage temperatures should be established. The influence of freeze-thaw cycles (a minimum of three cycles at two concentrations in triplicate) should be studied.

- The stability of the analyte in matrix at ambient temperature should be evaluated over a time period equal to the typical sample preparation, sample handling, and analytical run times.

- Reinjection reproducibility should be evaluated to determine if an analytical run could be reanalyzed in the case of instrument failure.

- The specificity of the assay methodology should be established using a minimum of six independent sources of the same matrix. For hyphenated mass spectrometry-based methods, however, testing six independent matrices for interference may not be important. In the case of LC-MS and LC-MS-MS-based procedures, matrix effects should be investigated to ensure that precision, selectivity, and sensitivity will not be compromised. Method selectivity should be evaluated during method development and throughout method validation and can continue throughout application of the method to actual study samples.

- Acceptance/rejection criteria for spiked, matrix-based calibration standards and validation QC samples should be based on the nominal (theoretical) concentration of analytes. Specific criteria can be set up in advance and achieved for accuracy and precision over the range of the standards, if so desired.

V. METHOD DEVELOPMENT: MICROBIOLOGICAL AND LIGAND-BINDING ASSAYS

Many of the bioanalytical validation parameters and principles discussed above are also applicable to microbiological and ligand-binding assays. However, these assays possess some unique characteristics that should be considered during method validation.

A. Selectivity Issues

As with chromatographic methods, microbiological and ligand-binding assays should be shown to be selective for the analyte. The following recommendations for dealing with two selectivity issues should be considered:

1. Interference From Substances Physiochemically Similar to the Analyte

- Cross-reactivity of metabolites, concomitant medications, or endogenous compounds should be evaluated individually and in combination with the analyte of interest.

- When possible, the immunoassay should be compared with a validated reference method (such as LC-MS) using incurred samples and predetermined criteria for agreement of accuracy of immunoassay and reference method.

- The dilutional linearity to the reference standard should be assessed using study (incurred) samples.

- Selectivity may be improved for some analytes by incorporation of separation steps prior to immunoassay.

2. Matrix effects Unrelated to the Analyte

- The standard curve in biological fluids should be compared with standard in buffer to detect matrix effects.

- Parallelism of diluted study samples should be evaluated with diluted standards to detect matrix effects.

- Nonspecific binding should be determined.

B. Quantification Issues

Microbiological and imunoassay standard curves are inherently nonlinear and, in general, more concentration points may be recommended to define the fit over the standard curve range than for chemical assays. In addition to their nonlinear characteristics, the response-error relationship for immunoassay standard curves is a nonconstant function of the mean response (heteroscadisticity). For these reasons, a minimum of six non-zero calibrator concentrations, run in duplicate, is recommended. The concentration-response relationship is most often fitted to a 4- or 5-parameter logistic model, although others may be used with suitable validation. The use of *anchoring points* in the asymptotic high- and low-concentration ends of the standard curve may improve the overall curve fit. Generally, these anchoring points will be at concentrations that are below the established LLOQ and above the established ULOQ. Whenever possible, calibrators should be prepared in the same matrix as the study samples or in an alternate matrix of equivalent performance. Both ULOQ and LLOQ should be defined by acceptable accuracy, precision, or confidence interval criteria based on the study requirements.

For all assays the key factor is the accuracy of the *reported results*. This accuracy can be improved by the use of replicate samples. In the case where replicate samples should be measured during the validation to improve accuracy, the same procedure should be followed as for unknown samples.

The following recommendations apply to quantification issues:

- If separation is used prior to assay for study samples but not for standards, it is important to establish recovery and use it in determining results. Possible approaches to assess efficiency

and reproducibility of recovery are (1) the use of radiolabeled tracer analyte (quantity too small to affect the assay), (2) the advance establishment of reproducible recovery, (3) the use of an internal standard that is not recognized by the antibody but can be measured by another technique.

- Key reagents, such as antibody, tracer, reference standard, and matrix should be characterized appropriately and stored under defined conditions.

- Assessments of analyte stability should be conducted in true study matrix (e.g., should not use a matrix stripped to remove endogenous interferences).

- Acceptance criteria: At least 67% (4 out of 6) of QC samples should be within 15% of their respective nominal value, 33% of the QC samples (not all replicates at the same concentration) may be outside 15% of nominal value. In certain situations, wider acceptance criteria may be justified.

- Assay reoptimization or validation may be important when there are changes in key reagents, as follows:

Labeled analyte (tracer)
- Binding should be reoptimized.
- Performance should be verified with standard curve and QCs.

Antibody
- Key cross-reactivities should be checked.
- Tracer experiments above should be repeated.

Matrix
- Tracer experiments above should be repeated.

Method development experiments should include a minimum of six runs conducted over several days, with at least four concentrations (LLOQ, low, medium, and high) analyzed in duplicate in each run.

VI. APPLICATION OF VALIDATED METHOD TO ROUTINE DRUG ANALYSIS

Assays of all samples of an analyte in a biological matrix should be completed within the time period for which stability data are available. In general, biological samples can be analyzed with a single determination without duplicate or replicate analysis if the assay method has acceptable variability as defined by validation data. This is true for procedures where precision and accuracy variabilities routinely fall within acceptable tolerance limits. For a difficult procedure with a labile analyte where high precision and accuracy specifications may be difficult to achieve, duplicate or even triplicate analyses can be performed for a better estimate of analyte.

A calibration curve should be generated for each analyte to assay samples in each analytical run and should be used to calculate the concentration of the analyte in the unknown samples in the run. The spiked samples can contain more than one analyte. An analytical run can consist of QC samples, calibration standards, and either (1) all the processed samples to be analyzed as one batch or (2) a batch composed of processed unknown samples of one or more volunteers in a study. The calibration (standard) curve should cover the expected unknown sample concentration range in addition to a calibrator sample at LLOQ. Estimation of concentration in unknown samples by extrapolation of standard curves below LLOQ or above the highest standard is not recommended. Instead, the standard curve should be redefined or samples with higher concentration should be diluted and reassayed. It is preferable to analyze all study samples from a subject in a single run.

Once the analytical method has been validated for routine use, its accuracy and precision should be monitored regularly to ensure that the method continues to perform satisfactorily. To achieve this objective, a number of QC samples prepared separately should be analyzed with processed test samples at intervals based on the total number of samples. The QC samples in duplicate at three concentrations (one near the LLOQ (i.e., #3 x LLOQ), one in midrange, and one close to the high end of the range) should be incorporated in each assay run. The number of QC samples (in multiples of three) will depend on the total number of samples in the run. The results of the QC samples provide the basis of accepting or rejecting the run. At least four of every six QC samples should be within "15% of their respective nominal value. Two of the six QC samples may be outside the "15% of their respective nominal value, but not both at the same concentration.

The following recommendations should be noted in applying a bioanalytical method to routine drug analysis:

- A matrix-based standard curve should consist of a minimum of six standard points, excluding blanks (either single or replicate), covering the entire range.

- Response Function: Typically, the same curve fitting, weighting, and goodness of fit determined during prestudy validation should be used for the standard curve within the study. Response function is determined by appropriate statistical tests based on the actual standard points during each run in the validation. Changes in the response function relationship between prestudy validation and routine run validation indicate potential problems.

- The QC samples should be used to accept or reject the run. These QC samples are matrix spiked with analyte.

- System suitability: Based on the analyte and technique, a specific SOP (or sample) should be identified to ensure optimum operation of the system used.

- Any required sample dilutions should use like matrix (e.g., human to human) obviating the need to incorporate actual within-study dilution matrix QC samples.

- Repeat Analysis: It is important to establish an SOP or guideline for repeat analysis and acceptance criteria. This SOP or guideline should explain the reasons for repeating sample analysis. Reasons for repeat analyses could include repeat analysis of clinical or preclinical samples for regulatory purposes, inconsistent replicate analysis, samples outside of the assay range, sample processing errors, equipment failure, poor chromatography, and inconsistent pharmacokinetic data. Reassays should be done in triplicate if sample volume allows. The rationale for the repeat analysis and the reporting of the repeat analysis should be clearly documented.

- Sample Data Reintegration: An SOP or guideline for sample data reintegration should be established. This SOP or guideline should explain the reasons for reintegration and how the reintegration is to be performed. The rationale for the reintegration should be clearly described and documented. Original and reintegration data should be reported.

Acceptance Criteria for the Run

The following acceptance criteria should be considered for accepting the analytical run:

- Standards and QC samples can be prepared from the same spiking stock solution, provided the solution stability and accuracy have been verified. A single source of matrix may also be used, provided selectivity has been verified.

- Standard curve samples, blanks, QCs, and study samples can be arranged as considered appropriate within the run.

- Placement of standards and QC samples within a run should be designed to detect assay drift over the run.

- Matrix-based standard calibration samples: 75%, or a minimum of six standards, when back-calculated (including ULOQ) should fall within ±15%, except for LLOQ, when it should be ±20% of the nominal value. Values falling outside these limits can be discarded, provided they do not change the established model.

- Acceptance criteria for accuracy and precision as outlined in section IV.F, "Specific Recommendation for Method Validation," should be provided for both the intra-day and intra-run experiment.

- Quality Control Samples: Quality control samples replicated (at least once) at a minimum of three concentrations (one within 3x of the LLOQ (low QC), one in the midrange (middle

QC), and one approaching the high end of the range (high QC)) should be incorporated into each run. The results of the QC samples provide the basis of accepting or rejecting the run. At least 67% (four out of six) of the QC samples should be within 15% of their respective nominal (theoretical) values; 33% of the QC samples (not all replicates at the same concentration) can be outside the ±15% of the nominal value. A confidence interval approach yielding comparable accuracy and precision is an appropriate alternative.

The minimum number of samples (in multiples of three) should be at least 5% of the number of unknown samples or six total QCs, whichever is greater.

- Samples involving multiple analytes should not be rejected based on the data from one analyte failing the acceptance criteria.

- The data from rejected runs need not be documented, but the fact that a run was rejected and the reason for failure should be recorded.

VII. DOCUMENTATION

The validity of an analytical method should be established and verified by laboratory studies, and documentation of successful completion of such studies should be provided in the assay validation report. General and specific SOPs and good record keeping are an essential part of a validated analytical method. The data generated for bioanalytical method establishment and the QCs should be documented and available for data audit and inspection. Documentation for submission to the Agency should include (1) summary information, (2) method development and establishment, (3) bioanalytical reports of the application of any methods to routine sample analysis, and (4) other information applicable to method development and establishment and/or to routine sample analysis.

A. Summary Information

Summary information should include:

- Summary table of validation reports, including analytical method validation, partial revalidation, and cross-validation reports. The table should be in chronological sequence, and include assay method identification code, type of assay, and the reason for the new method or additional validation (e.g., to lower the limit of quantitation).

- Summary table with a list, by protocol, of assay methods used. The protocol number, protocol title, assay type, assay method identification code, and bioanalytic report code should be provided.

- A summary table allowing cross-referencing of multiple identification codes should be provided (e.g., when an assay has different codes for the assay method, validation reports,

and bioanalytical reports, especially when the sponsor and a contract laboratory assign different codes).

B. Documentation for Method Establishment

Documentation for method development and establishment should include:

- An operational description of the analytical method
- Evidence of purity and identity of drug standards, metabolite standards, and internal standards used in validation experiments
- A description of stability studies and supporting data
- A description of experiments conducted to determine accuracy, precision, recovery, selectivity, limit of quantification, calibration curve (equations and weighting functions used, if any), and relevant data obtained from these studies
- Documentation of intra- and inter-assay precision and accuracy
- In NDA submissions, information about cross-validation study data, if applicable
- Legible annotated chromatograms or mass spectrograms, if applicable
- Any deviations from SOPs, protocols, or GLPs (if applicable), and justifications for deviations

C. Application to Routine Drug Analysis

Documentation of the application of validated bioanalytical methods to routine drug analysis should include:

- Evidence of purity and identity of drug standards, metabolite standards, and internal standards used during routine analyses
- Summary tables containing information on sample processing and storage. Tables should include sample identification, collection dates, storage prior to shipment, information on shipment batch, and storage prior to analysis. Information should include dates, times, sample condition, and any deviation from protocols.
- Summary tables of analytical runs of clinical or preclinical samples. Information should include assay run identification, date and time of analysis, assay method, analysts, start and

stop times, duration, significant equipment and material changes, and any potential issues or deviation from the established method.

- Equations used for back-calculation of results

- Tables of calibration curve data used in analyzing samples and calibration curve summary data

- Summary information on intra- and inter-assay values of QC samples and data on intra- and inter-assay accuracy and precision from calibration curves and QC samples used for accepting the analytical run. QC graphs and trend analyses in addition to raw data and summary statistics are encouraged.

- Data tables from analytical runs of clinical or preclinical samples. Tables should include assay run identification, sample identification, raw data and back-calculated results, integration codes, and/or other reporting codes.

- Complete serial chromatograms from 5**B**20% of subjects, with standards and QC samples from those analytical runs. For pivotal bioequivalence studies for marketing, chromatograms from 20% of serially selected subjects should be included. In other studies, chromatograms from 5% of randomly selected subjects in each study should be included. Subjects whose chromatograms are to be submitted should be defined prior to the analysis of any clinical samples.

- Reasons for missing samples

- Documentation for repeat analyses. Documentation should include the initial and repeat analysis results, the reported result, assay run identification, the reason for the repeat analysis, the requestor of the repeat analysis, and the manager authorizing reanalysis. Repeat analysis of a clinical or preclinical sample should be performed only under a predefined SOP.

- Documentation for reintegrated data. Documentation should include the initial and repeat integration results, the method used for reintegration, the reported result, assay run identification, the reason for the reintegration, the requestor of the reintegration, and the manager authorizing reintegration. Reintegration of a clinical or preclinical sample should be performed only under a predefined SOP.

- Deviations from the analysis protocol or SOP, with reasons and justifications for the deviations

D. Other Information

Other information applicable to both method development and establishment and/or to routine sample analysis could include:

- Lists of abbreviations and any additional codes used, including sample condition codes, integration codes, and reporting codes

- Reference lists and legible copies of any references

- SOPs or protocols covering the following areas:

    - Calibration standard acceptance or rejection criteria
    - Calibration curve acceptance or rejection criteria
    - Quality control sample and assay run acceptance or rejection criteria
    - Acceptance criteria for reported values when all unknown samples are assayed in duplicate
    - Sample code designations, including clinical or preclinical sample codes and bioassay sample code
    - Assignment of clinical or preclinical samples to assay batches
    - Sample collection, processing, and storage
    - Repeat analyses of samples
    - Reintegration of samples

GLOSSARY

Accuracy: The degree of closeness of the determined value to the nominal or known true value under prescribed conditions. This is sometimes termed *trueness*.

Analyte: A specific chemical moiety being measured, which can be intact drug, biomolecule or its derivative, metabolite, and/or degradation product in a biologic matrix.

Analytical run (or batch): A complete set of analytical and study samples with appropriate number of standards and QCs for their validation. Several runs (or batches) may be completed in one day, or one run (or batch) may take several days to complete.

Biological matrix: A discrete material of biological origin that can be sampled and processed in a reproducible manner. Examples are blood, serum, plasma, urine, feces, saliva, sputum, and various discrete tissues.

Calibration standard: A biological matrix to which a known amount of analyte has been added or *spiked*. Calibration standards are used to construct calibration curves from which the concentrations of analytes in QCs and in unknown study samples are determined.

Internal standard: Test compound(s) (e.g. structurally similar analog, stable labeled compound) added to both calibration standards and samples at known and constant concentration to facilitate quantification of the target analyte(s).

Limit of detection (LOD): The lowest concentration of an analyte that the bioanalytical procedure can reliably differentiate from background noise.

Lower limit of quantification (LLOQ): The lowest amount of an analyte in a sample that can be quantitatively determined with suitable precision and accuracy.

Matrix effect: The direct or indirect alteration or interference in response due to the presence of unintended analytes (for analysis) or other interfering substances in the sample.

Method: A comprehensive description of all procedures used in sample analysis.

Precision: The closeness of agreement (*degree of scatter*) between a series of measurements obtained from multiple sampling of the same homogenous sample under the prescribed conditions.

Processed: The final extract (prior to instrumental analysis) of a sample that has been subjected to various manipulations (e.g., extraction, dilution, concentration).

Quantification range: The range of concentration, including ULOQ and LLOQ, that can be reliably and reproducibly quantified with accuracy and precision through the use of a concentration-response relationship.

Recovery: The extraction efficiency of an analytical process, reported as a percentage of the known amount of an analyte carried through the sample extraction and processing steps of the method.

Reproducibility: The precision between two laboratories. It also represents precision of the method under the same operating conditions over a short period of time.

Sample: A generic term encompassing controls, blanks, unknowns, and processed samples, as described below:

> **Blank:** A sample of a biological matrix to which no analytes have been added that is used to assess the specificity of the bioanalytical method.
>
> **Quality control sample (QC):** A spiked sample used to monitor the performance of a bioanalytical method and to assess the integrity and validity of the results of the unknown samples analyzed in an individual batch.
>
> **Unknown:** A biological sample that is the subject of the analysis.

Selectivity: The ability of the bioanalytical method to measure and differentiate the analytes in the presence of components that may be expected to be present. These could include metabolites, impurities, degradants, or matrix components.

Stability: The chemical stability of an analyte in a given matrix under specific conditions for given time intervals.

Standard curve: The relationship between the experimental response value and the analytical concentration (also called a *calibration curve*).

System suitability: Determination of instrument performance (e.g., sensitivity and chromatographic retention) by analysis of a reference standard prior to running the analytical batch.

Upper limit of quantification (ULOQ): The highest amount of an analyte in a sample that can be quantitatively determined with precision and accuracy.

Validation:

> **Full validation:** Establishment of all validation parameters to apply to sample analysis for the bioanalytical method for each analyte.

Partial validation: Modification of validated bioanalytical methods that do not necessarily call for full revalidation.

Cross-validation: Comparison validation parameters of two bioanalytical methods.