ANSI/ASB Standard 036, First Edition
2019

Standard Practices for Method Validation in Forensic Toxicology

Standard Practices for Method Validation in Forensic Toxicology

ASB Approved March 2019

ANSI Approved November 2019

410 North 21st Street
Colorado Springs, CO 80904

Foreword

Validation is the process of performing a set of experiments to establish objective evidence that a method is fit-for-purpose and to identify the method's limitations under normal operating conditions. This standard was developed to provide the minimum requirements for validating analytical methods in forensic toxicology laboratories.

This document was revised, prepared, and finalized as a standard by the Toxicology Consensus Body of the AAFS ASB. It was originally drafted by the Scientific Working Group on Forensic Toxicology (SWGTOX). When SWGTOX disbanded in 2014, it passed ownership of all of its documents to the Toxicology Subcommittee of the Organization of Scientific Area Committees (OSAC), who in turn updated and approved the draft document.

All hyperlinks and web addresses shown in this document are current as of the publication date of this standard.

Keywords: *validation, forensic toxicology*

Table of Contents

Figures

Tables

Standard Practices for Method Validation in Forensic Toxicology

1 Scope

This document delineates minimum standards of practice for validating analytical methods used in the field of forensic toxicology that target specific analytes or analyte classes. Specifically, it is intended for the subdisciplines of postmortem forensic toxicology, human performance toxicology (e.g., drug-facilitated crimes and driving-under-the-influence of alcohol or drugs), non-regulated employment drug testing, court-ordered toxicology (e.g., probation and parole, drug courts, child services), and general forensic toxicology (non-lethal poisonings or intoxications). This document is not intended to address method validation in the discipline of breath alcohol testing. The fundamental reason for performing method validation is to ensure confidence and reliability in forensic toxicological test results by demonstrating the method is fit for its intended use.

2 Normative References

There are no normative references. Annex F, Bibliography, contains informative references.

3 Terms and Definitions

For purposes of this document, the following definitions and acronyms apply.

3.1
bias
An estimate of systematic measurement error, calculated as the difference between the mean of several measurements under identical conditions, to a known "true" value. It is often reported as a percent difference.

3.2
biological fluid
Any liquid biological specimen that is typically pipetted for analysis (e.g., blood, urine, bile, serum, vitreous humor, oral fluid).

3.3
blank matrix sample
A biological fluid or tissue (or synthetic substitute) without target analyte or internal standard.

3.4
calibration model
The mathematical model that demonstrates the relationship between the concentration of analyte and the corresponding instrument response.

3.5
carryover
The appearance of unintended analyte signal in samples after the analysis of a positive sample.

3.6
decision point
An administratively defined cutoff or concentration that is at or above the method's limit of detection or limit of quantitation and is used to discriminate between positive and negative results.

3.7
fortified matrix sample
A blank matrix sample spiked with target analyte and/or internal standard using reference materials.

3.8
interferences
Non-targeted analytes (i.e., matrix components, other drugs and metabolites, internal standard, impurities) which may impact the ability to detect, identify, or quantitate a targeted analyte.

3.9
ionization suppression/enhancement
Direct or indirect alteration or interference in the instrument response due to the presence of co-eluting compounds.

3.10
limit of detection
LOD
An estimate of the lowest concentration of an analyte in a sample that can be reliably differentiated from blank matrix and identified by the analytical method.

3.11
lower limit of quantitation
LLOQ
An estimate of the lowest concentration of an analyte in a sample that can be reliably measured with acceptable bias and precision.

3.12
precision
The measure of the closeness of agreement between a series of measurements obtained from multiple samplings of the same homogenous sample. It is expressed numerically as the coefficient of variation (%CV).

3.13
reference material
Material, sufficiently homogenous and stable with reference to specified properties, which have been established to be fit for its intended use in a measurement or in examination of nominal properties.

3.14
qualitative confirmation/identification method
An assay designed to be more specific for targeted analytes and based on a different chemical principle than a screening method.

3.15
quantitative method
An assay designed to measure the concentration of an analyte within a sample.

3.16
screening method
An assay designed to rule out the presence of analytes or to suggest their presence; thereby indicating further testing may be warranted.

3.17
stability
An analyte's resistance to chemical change in a matrix under specific conditions for given time intervals.

3.18
tissues
Any solid biological specimen that is generally weighed for analysis (e.g., brain, liver, muscle, hair, bone, meconium).

3.19
working range
The range of concentrations that can be adequately determined by an instrument, where the instrument provides a useful signal that can be related to the concentration of the analyte.

4 When to Validate Methods

Methods shall be validated to verify a method's performance parameters are fit for use for a particular analysis. Common examples include:

a) new analytical method;

b) modifications of an established analytical method to improve performance or extend its use beyond that for which it was originally validated (e.g., addition of new compounds to the method's scope);

c) to demonstrate equivalence between an established method/instrument and a new method/instrument;

d) existing analytical methods that do not currently meet the requirements of this document.

The parameters to be evaluated for validation of methods will depend upon the circumstances in which the method is to be used. Likewise, it is recognized that after validation has occurred, methods may be revised. The extent and frequency of revalidation of previously validated methods will depend upon the nature of the intended changes or laboratory policy. See Section 10 for further guidance on revalidation of previously validated methods.

Laboratories using methods that were validated prior to the promulgation of this document shall demonstrate and document that those methods are fit-for-purpose under this standard. These methods will likely have sufficient historical calibration and control data, as well as previously analyzed casework sample results, that can be used to address a number of the required validation parameters. In the absence of sufficient data to fulfill these minimum standards, appropriate studies shall be conducted to ensure compliance with this document.

5 Method Development and Optimization

5.1 General

For purposes of this document, method development will be considered in two phases: 1) instrumental and data acquisition/processing parameters and 2) sample preparation. It is essential that validation is conducted with the same analytical conditions and techniques as the final developed method.

The principles of good laboratory practice and record keeping shall be applied to the concepts of this document. This includes documentation of parameters that were evaluated during method development, yet did not provide acceptable results.

5.2 Development and Optimization of Instrumental and Data Processing Parameters

Instrumental and data processing parameters are defined and optimized through analysis of reference materials of the analyte(s) of interest to achieve the required performance of the instrument.

5.3 Development and Optimization of Sample Preparation Techniques

The sample preparation technique shall be evaluated and optimized using reference materials of the analyte(s) of interest. The primary goal is to demonstrate that the sample preparation steps allow for adequate extraction, detection, identification, and/or quantitation of the analyte(s). Sample preparation shall be evaluated with fortified matrix samples.

6 Establishing a Validation Plan

A validation plan shall be in place prior to starting any validation experiments. The validation plan is separate from a laboratory's standard operating procedure for method validation. The plan shall include the instrumental method(s) and sample preparation technique(s) to be used for a specific method. Further, it shall document the validation requirements of the method, as well as the limits of the method that will allow it to be fit for use. The validation plan provides direction for the experiments that will be performed and acceptance criteria for each parameter. Annex A and Annex B provide examples of validation plans.

7 Required Validation Parameters Based on Scope of the Method

7.1 The scope of forensic toxicology methods is typically categorized as screening methods, qualitative confirmation/identification methods, or quantitative methods. As such the following validation parameters shall be evaluated.

7.2 Screening (Immunoassay-based):

a) limit of detection;

b) precision (at the decision point);

c) processed sample stability (if applicable).

7.3 Screening (all other methods targeting specific analytes or analyte classes):

a) interference studies;

b) limit of detection;

c) ionization suppression/enhancement [for applicable techniques, such as liquid chromatography/mass spectrometry (LC/MS)];

d) processed sample stability (if applicable).

7.4 Qualitative confirmation/identification:

a) carryover;

b) interference studies;

c) ionization suppression/enhancement (for applicable techniques, such as LC/MS);

d) limit of detection;

e) processed sample stability (if applicable).

7.5 Quantitative analysis:

a) bias;

b) calibration model;

c) carryover;

d) interference studies;

e) ionization suppression/enhancement (for applicable techniques, such as LC/MS);

f) limit of detection;

g) limit of quantitation;

h) precision;

i) dilution integrity (if applicable);

j) processed sample stability (if applicable).

8 Specific Requirements for Conducting Method Validation Experiments

8.1 General

All validation experiments shall be conducted using fortified samples for each matrix type which the method is intended, unless otherwise noted. For example, a method with a scope to analyze both postmortem blood and urine specimens shall include complete validation experiments using fortified blank postmortem blood samples and complete validation experiments using fortified blank urine samples. In some instances, it may be more appropriate to analyze previously characterized human samples instead of fortified samples for selected method validation studies (e.g., interference studies).

Validation studies shall be conducted in a manner similar to casework. This may include conducting validation studies on different days, by different analysts, on all identical instruments to be utilized for the assay, and ensuring that the instruments meet the same daily performance requirements as for casework.

Fortified matrix samples should be prepared from reference materials that are from a different source (e.g., supplier or lot number) than used to prepare calibration samples. In instances where the same source shall be utilized, separate weighings or solutions shall be used to prepare these samples.

The following requirements are the minimum for assessing the listed validation parameters in forensic toxicology methods. They are listed alphabetically and not necessarily in procedural order. Some of the validation experiments are demonstrated in Annex A and Annex B. Annexes C, D, and E provide guidance on how to efficiently perform validation experiments.

8.2 Bias and Precision[1]

8.2.1 Bias

Bias studies shall be carried out for all quantitative methods. These can be conducted concurrently with precision studies.

Bias shall be measured in pooled fortified matrix samples using a minimum of three separate samples per concentration at three different concentration pools (low, medium and high[2]) over five different runs.[3] The bias shall be calculated for each concentration using the following formula:

$$\textbf{\textit{Bias}}\ (\%)\ \textbf{\textit{at Concentration}}_x = \left[\frac{\textbf{\textit{Grand Mean of Calculated Concentration}}_x - \textbf{\textit{Nominal Concentration}}_x}{\textbf{\textit{Nominal Concentration}}_x}\right] \times \mathbf{100}$$

[1] An accurate measurement is one with acceptable bias and precision.

[2] For purposes of this document, low concentrations shall be no more than approximately 3 times the lowest end of the working range of the method and high concentrations shall be within approximately 80% (or more) of the highest end of the working range of the method, unless otherwise noted. Medium concentrations shall be near the midpoint of the low and high concentrations.

[3] In some instances, analyte instability may preclude the ability to use concentration pools of fortified samples (e.g., cocaine in unpreserved whole blood). In these instances, a laboratory may resort to fortifying different samples with each run.

The maximum acceptable bias shall be ±20% at each concentration. For some analyses where less bias is required (e.g., ethanol), a bias of ±10% or better should be expected. It is recommended that the same data used in bias studies also be used for precision calculations.

8.2.2 Precision

8.2.2.1 General

Precision studies shall be carried out for all quantitative methods, as well as at the decision point for immunoassays. These studies can be carried out concurrently with bias studies, if required in the validation plan.

Precision is expressed as the coefficient of variation (% CV). The mean and standard deviation (*std dev)* of the response is calculated for each concentration to determine the % CV.

$$\%\boldsymbol{CV} = \frac{\boldsymbol{std\ dev}}{\boldsymbol{mean\ response}}\ \boldsymbol{x}\ \mathbf{100}$$

8.2.2.2 Precision of Immunoassays at Decision Point

At a minimum, precision at the immunoassay's decision point (i.e. cutoff concentration) shall be assessed for the target analyte using three separate samples from three different concentration pools over five different runs:

a) generally no more than 50% below the decision point,

b) at the decision point, and

c) generally no more than 100% above the decision point.

The intent of this requirement is to ensure that there is evidence that the assay separates presumptive positive and negative samples at the decision point.

For Enzyme Linked Immunosorbent Assays (ELISA), the difference between the absorbance of a negative sample (B_0) and the absorbance of a specimen (B) should be used as a percentage: [B/B_0] x 100 and not an absolute value. Note that B_0 is established for each run.

For liquid reagent assays (e.g., enzyme multiplied immunoassay technique [EMIT], cloned enzyme donor immunoassay [CEDIA®[4]], etc.), the absorbance value may be directly used.

a) Run each concentration 3 times in five separate runs. Additionally, for ELISA assays, at least one negative sample per run shall be analyzed to establish B_0.

b) Calculate the grand mean (n=15) and related grand standard deviation for each concentration pool.

[4] This term is used as an example only, and does not constitute an endorsement of this product by the AAFS Standards Board.

c) The %CV shall not exceed 20% at each concentration using all 15 sample results per concentration.

d) The grand mean plus or minus two standard deviations of the low and high concentration pools shall not overlap with the mean of the decision point.

8.2.2.3 Precision of Quantitative Procedures

8.2.2.3.1 General

For quantitative procedures, two different types of precision studies shall be assessed during method validation: within-run precision and between-run precision. At a minimum, precision shall be assessed using triplicate analyses per concentration pool (low, medium and high) over five different runs. The different runs used to evaluate precision may be performed within the same day, provided a different calibration curve is used for each run.

The % CV shall not exceed 20% at each concentration. It is noted that certain analytical methods (e.g., blood alcohol analysis) should require a much lower coefficient of variation (≤10% CV).

The largest calculated within-run and between-run % CV for each concentration shall be used to assess precision acceptability.

8.2.2.3.2 Within-Run Precision Calculations

Within-run precisions are calculated for each concentration separately for each of the five runs. Within-run precision may be calculated using the data from each run's triplicate analyses at each concentration as:

$$Within-run\ CV(\%) = \frac{std\ dev of\ a\ single\ run\ of\ samples}{mean\ calculated\ value\ of\ a\ single\ run\ of\ samples} \times 100$$

8.2.2.3.3 Between-Run Precision Calculations

Between-run precision is calculated for each concentration over the five runs. This may be done by using the combined data from all replicates of each concentration as:

$$Between-Run\ CV(\%) = \frac{std\ dev\ of\ all\ observations\ for\ each\ concentration}{grand\ mean\ for\ each\ concentration} \times 100$$

8.2.2.3.4 One-Way Analysis of Variance (ANOVA) Approach to Calculate Combined Within-Run and Between-Run Precision

Both within-run and between-run precisions may be calculated using the one-way ANOVA approach with the varied factor (run number) as the grouping variable. The ANOVA calculations can be easily performed using a spreadsheet or a statistical software program.

Using this approach, within-run precisions are calculated for each concentration as:

$$Within-run\ CV(\%) = \left[\frac{\sqrt{MS_{wg}}}{grand\ mean\ for\ each\ concentration}\right] \times 100$$

where MS_{wg} is the mean square within groups obtained from the ANOVA table.

Likewise, between-run precisions are calculated as:

$$Between - Run\ CV(\%) = \left[\frac{\sqrt{\frac{MS_{bg} + (n-1) * MS_{wg}}{n}}}{grand\ mean\ for\ each\ concentration}\right] \times 100$$

where MS_{bg} is the mean square between groups obtained from the ANOVA table and *n* is the number of observations in each group (e.g., *n*=3 if doing triplicate analyses). Annex A provides an example of how the ANOVA approach may be used to calculate within-run and between-run precision.

8.3 Calibration Model

The calibration model shall be determined for all quantitative methods. This is accomplished by first determining the working range. Within this range, there will be a correlation between signal response (e.g., peak area ratio of analyte and internal standard) and analyte concentration in the sample. The calibration model is the mathematical model that describes this correlation. The choice of an appropriate model (i.e., linear or quadratic) is necessary for accurate and reliable quantitative results.

Calibrator samples are analyzed to establish the calibration model. The use of matrix-matched calibrator samples is encouraged, but not required. A single source of blank matrix (per matrix type) may be used when experimentally establishing the appropriate calibration model for a method. Regardless of the matrix used to prepare calibrator samples, a laboratory shall demonstrate acceptable bias and precision with control samples prepared in all matrices intended to be analyzed by the method (see Section 7.1). For example, blood alcohol methods may demonstrate acceptable bias and precision in whole blood controls using aqueous calibrator samples. Likewise, acceptable bias and precision may be demonstrated using calibrator samples prepared in whole blood but used to quantitate analytes in different matrices (e.g., postmortem tissues, serum, urine).

The calibrator samples shall span the range of concentrations expected in day-to-day operations. At least six different non-zero concentrations shall be used to evaluate the calibration model. The concentrations shall be appropriately spaced across the calibration range to establish the most appropriate calibration model. A minimum of five replicates per concentration shall be used. The replicates to establish the calibration model may be in the same or in separate runs. All data points from the five replicates shall be plotted together (using a statistical software package) to establish the calibration model. The origin shall not be included as a calibration point.

The simple linear regression model using the least squares method is the most often used calibration model. However, this model is only applicable when there is constant variance over the entire concentration range (homoscedasticity). When there is a notable difference between variances at the lowest and highest concentrations, a weighted least squares model or other appropriate non-linear model should be applied.[5] This is generally the case when the concentration

[5] In general, non-linear models may require additional calibrators to accurately characterize the curve.

range exceeds one order of magnitude. Ultimately, the simplest calibration model that best fits the concentration-response relationship should be used.

A calibration model shall not be evaluated simply via its correlation coefficient (*r*). Instead, a calibration model should be visually evaluated using residual plots. These allow one to check for outliers that may be eliminated if found to be statistically significant (e.g., outside ±3 standard deviations). Further, residual plots allow one to determine if the variances appear to be equal across the calibration range with a similar degree of scatter at each concentration. They also give an indication if the chosen model adequately fits the data. For example, random distribution of individual residuals around the zero line suggests that a linear model is appropriate.

Finally, there are other appropriate alternatives to evaluate calibration models (i.e., ANOVA lack-of-fit test for unweighted linear models, checking for significance of the second order term in quadratic models, assessment of coefficient of determination (r^2) for linear models).

If a linear calibration model has been established, fewer calibration samples (i.e., fewer levels or single/fewer replicates) may be used for routine analysis. However, if fewer calibration samples are chosen, the same calibrators (e.g., number, replicates, and concentration levels) shall be used to construct the calibration curves for the bias and precision studies. Further the calibration data shall include the lowest and highest calibration levels used to establish the model, as well as include no fewer than four non-zero calibration points.

Additionally, once the calibration model is established for a validated method, it shall not be arbitrarily changed to achieve acceptable quality control results during a given analytical run. For example, one shall not switch from an unweighted linear model to a weighted linear model in order to adjust for changes in instrument performance without revalidation of the impact on the change.

8.4 Carryover

Analyte carryover into a subsequent sample may lead to an inaccurate qualitative or quantitative result when using instrumental methods. Carryover shall be evaluated during method validation intended for confirmation and/or quantitation, unless a laboratory is constantly addressing carryover in their quality assurance (QA)/quality control (QC) practices.

To evaluate carryover as part of method validation, blank matrix samples are analyzed immediately after a high concentration sample or reference material. The highest fortified concentration at which no analyte carryover is observed (above the method's LOD) in the blank matrix sample is determined to be the concentration at which the method is free from carryover. This carryover concentration for each analyte in the method shall be confirmed using triplicate analyses. It is acceptable to limit the carryover study to the highest point of your calibration curve in quantitative assays.

The analytical procedure should be modified to remove any carryover. In cases when it is not possible to eliminate the carryover, the SOP shall address how carryover will be managed.

8.5 Interference Studies

8.5.1 General

Interfering substances from common sources shall be evaluated in all screening (except immunoassays), qualitative identification, and quantitative methods.

8.5.2 Evaluating Matrix Interferences

Whenever possible, blank matrix samples from a minimum of ten different sources without the addition of an internal standard (when used in the method) shall be analyzed to demonstrate the absence of common interferences from the matrix. While this approach may detect the more common matrix interferences, it is recognized that less common interferences may not be detected.

8.5.3 Evaluating Interferences from Stable-Isotope Internal Standards

For methods employing stable isotope internal standards, the isotopically-labeled compounds may contain the non-labeled compound as an impurity. Additionally, the mass spectra of the labeled analogs may contain fragment ions with the same mass-to-charge ratios as the significant ions of the target analyte. In both instances, analyte identification or quantitation could be impacted.

Stable-isotope internal standard interferences shall be assessed by analyzing a single blank matrix sample (per matrix type) fortified with the internal standard(s) and monitoring the signal of the analyte(s) of interest. Interferences below the LOD of the assay may be insignificant depending on the laboratory's mission.

Likewise, a single blank matrix sample (per matrix type) fortified with the analyte(s) at a concentration near the upper limit of the calibration range shall be analyzed without internal standard to evaluate whether relevant amounts of the unlabeled analyte ions appear as isotopically-labeled compound fragments which could impact quantitation.

8.5.4 Evaluating Interferences from Other Commonly Encountered Analytes

For all methods other than immunoassays, it is necessary to evaluate other analytes which may be expected to be present in case samples for their potential to interfere with the method's analytes. For example, a method developed to analyze blood for cocaine shall evaluate whether other common drugs of abuse, metabolites, and structurally-similar compounds interfere with the assay. Likewise, a headspace gas chromatograph-flame ionization detection (GC-FID) method developed for ethanol shall evaluate whether other common volatile organic compounds interfere with the assay.

This evaluation shall be accomplished by analyzing fortified matrix samples, previously analyzed case samples, or neat reference materials of the potential interference(s) at high therapeutic or lethal concentrations, depending on the analyte, the matrix, and the laboratory's mission. The most common drugs/metabolites encountered in the laboratory shall be included in the evaluation together with other common drugs within the classification, where appropriate.

8.6 Ionization Suppression/Enhancement

8.6.1 General

The enhancement or suppression of analyte ionization resulting from the presence of co-eluting compounds is a phenomenon commonly encountered in LC/MS applications.

When average suppression/enhancement of the analyte's target ion (or ion transition and qualifying ions) exceeds ±25% or the % CV of the suppression/enhancement exceeds 20%, a laboratory shall demonstrate that there is no impact on other critical validation parameters. For example, suppression/enhancement of ionization is most likely to impact the LOD of a qualitative

method. Likewise, the LOD and LLOQ may be affected by ionization suppression/enhancement in quantitative methods. The influence on the above parameters shall be assessed by at least tripling the number of different sources of blank matrices used in their evaluation. For example, if the average suppression/enhancement exceeds ±25%, the LOD determination shall be performed in at least nine unique sources of blank matrices.

Laboratories shall also assess the impact of ionization suppression/enhancement on the method's internal standards.

Ionization suppression/enhancement shall be evaluated using either of the approaches that follow.

8.6.2 Post-column Infusion to Assess Ionization Suppression/Enhancement

This approach provides information on retention times where ionization suppression/enhancement occurs. It is useful during method development, but may also be used to assess the amount of ionization suppression/enhancement for LC/MS based confirmation methods. Solutions at both low and high concentrations of the analyte are individually infused with a syringe pump into the eluent from the column via a post-column "T"-connector and a constant baseline signal for the analyte of interest is monitored. Whenever possible, a minimum of ten different processed blank matrix samples (per matrix type) that are representative of the quality of samples typically encountered in casework are injected into the LC/MS during infusion of the solutions.[6] If there is any considerable suppression/enhancement (>25%) of the infused analyte signal at the retention time of the analyte, then modification of the chromatographic system or the sample preparation may be required to minimize the effect of ionization suppression/enhancement.

8.6.3 Post-Extraction Addition to Assess Ionization Suppression/Enhancement

This approach yields a quantitative estimation of ionization suppression/enhancement. It is useful for assessing the amount of ionization suppression/enhancement for LC/MS based methods. Two different sets of samples shall be prepared and the analyte peak areas of neat standards shall be compared to matrix samples fortified with neat standards after extraction or processing.

Set one consists of neat standards prepared at two concentrations – one low and one high. Each of these neat standards shall be injected a minimum of six times to establish a mean peak area for each concentration.

Set two should consist of a minimum of ten different matrix sources (per matrix type).[7] Each matrix source shall be extracted in duplicate. After the extraction is complete, each matrix sample shall be reconstituted/fortified with either the low or high concentration neat standard.

[6] Additional matrix samples may be needed in postmortem toxicology given the variety of sample conditions typically encountered in this work.
[7] Additional matrix samples may be needed in postmortem toxicology given the variety of sample conditions typically encountered in this work.

The average area of each set ($\overline{X}$) shall be used to estimate the suppression/ enhancement effect at each concentration as follows:

$$Ionization\ suppression\ or\ enhancement\ (\%) = \left(\frac{\overline{X}_{Area\ of\ Set\ 2}}{\overline{X}_{Area\ of\ Set\ 1}} - 1 \right) \times 100$$

Two ionization suppression/enhancement percentages shall be established – one at the low concentration and one at the high concentration.

8.7 Limit of Detection

8.7.1 General

LOD studies shall be carried out for all methods. There are a number of different approaches for determining the LOD. Select the approach that provides the most reasonable estimation of the LOD given the analytical instrumentation (or lack thereof) utilized in the method.

A method's LOD incorporates instrumental performance, as well as the sample matrix and inherent procedural limitations. Therefore, the LOD shall be assessed over multiple runs using fortified matrix samples from at least three different sources of blank matrix, unless otherwise indicated. For example, if the assay is to be used for postmortem blood samples, three independent representative postmortem blood sources are needed.

The LOD shall be determined by one of the following approaches.

8.7.2 Estimating LOD for Immunoassays

The laboratory may assign the decision point (i.e., cutoff concentration) as the LOD for immunoassays. Most of these assays are known to cross-react with numerous drugs (e.g., benzodiazepines, opiates, amphetamines) and metabolites. When a laboratory declares to their customers that they are able to detect specific analytes demonstating low cross-reactivity (less than or equal to the target analyte) using the immunoassay, they shall verify their ability to reliably detect these compounds.

As an example, a benzodiazepine immunoassay targeted for oxazepam typically has low cross-reactivities to many other benzodiazepines. If a laboratory uses the decision point (i.e. cut-off concentration) determined by the manufacturer, the laboratory shall mathematically estimate the LOD concentration for any benzodiazepines that cross-react at less than 100% *and* that they declare to their customers they are able to detect with the immunoassay. For example, if lorazepam cross-reacts at 50%, it would have an equivalent decision point concentration of 100 ng/mL in an oxazepam-targeted assay with a decision point of 50 ng/mL. In contrast, if another benzodiazepine (e.g., alprazolam) has greater cross-reactivity than the target drug, oxazepam, there is no requirement to mathematically estimate the LOD concentration for that drug *provided that the decision point for the drug (alprazolam) is not lower than the decision point (50 ng/mL) for the target analyte (oxazepam).*

In contrast, if a laboratory uses a different decision point than that recommended by the manufacturer of an immunoassay kit, the laboratory shall experimentally determine the LOD of the other benzodiazepines in the panel with cross-reactivities at or below that of the target analyte.

This evaluation may require an adjustment or reevaluation of the decision point or the target compound depending on the needs and mission of the laboratory. The experiments performed to assess precison (Section 8.2.2.2) shall be used to demonstrate a suitable LOD has been determined. The replicates for analytes with cross-reactivities at or below the target analyte (n=15) must show an average that indicates a "positive" response compared to the target analyte.

A single source of blank matrix (per matrix type) may be used to estimate the LOD of immunoassays.

8.7.3 Estimating LOD for a Non-Instrumental Method

This approach is most often used when screening for the presence or absence of a specified analyte or class of analytes (e.g., color tests). To estimate the LOD for a non-instrumental method, at least three different blank matrix sources fortified with decreasing concentrations of analyte shall be analyzed over a minimum of three runs. Multiple analysts should be involved in estimating the LOD using this approach, due to greater subjectivity in these methods. The lowest concentration of analyte that yields a positive result on all runs and observed by all participating analysts shall be considered the LOD.

8.7.4 Using the Lowest Non-Zero Calibrator as the LOD

This technique is useful for quantitative methods. In some instances, it may be sufficient to define the LOD as the value of the lowest non-zero calibrator. For each matrix type, at least three different blank matrix sources shall be fortified with the analyte at the concentration of the lowest calibrator and analyzed over a minimum of three runs to demonstrate that all detection and identification criteria are met. If desired, it is acceptable to use the same calibrator replicates used to establish the calibration model (Section 8.3) for this approach, but additional samples/replicates may be needed to meet the minimum of nine data points including at least three sources per matrix type.

8.7.5 Using the Decision Point Concentration as the LOD

This technique is useful for qualitative and quantitative methods. In some instances, it may be sufficient to define the LOD as the value of an administratively-defined decision point. For example, a laboratory may choose to define a method's LOD for ethanol as 0.02 g/dL for blood based on the laboratory's administratively defined decision point for reporting this analyte, even though a lower LOD is analytically achievable. For each matrix type, at least three different blank matrix sources shall be fortified with the analyte at the concentration of the decision point and analyzed over a minimum of three runs to demonstrate that all detection and identification criteria are met.

8.7.6 Estimating LOD Using Background Noise

8.7.6.1 General

These approaches for determining LOD are only useful for instrumental methods that demonstrate background noise. A minimum of three different blank sources per matrix type shall be used.

8.7.6.2 Estimating LOD Using Reference Materials

Three (or more) sources of blank matrix samples (per matrix type) fortified at decreasing concentrations shall be analyzed in duplicate (two separate replicates) for at least three runs. The LOD shall be the lowest concentration that 1) yields a reproducible instrument response greater

than or equal to 3.3 times[8] the noise level of the background signal in an area around the analyte peak *and* 2) achieves acceptable predefined detection and identification criteria (e.g., retention time, peak shape, mass spectral ion ratios) for all replicates at that concentration.

While it may be possible to visually assess the signal-to-noise ratio, such an approach is subjective. Therefore, the signal-to-noise ratio shall be manually calculated *or* determined through use of instrumentation software. If manually calculated, the signal is defined as the height response of the analyte peak and the noise is defined as the amplitude between the highest and lowest point of the baseline in an area around the analyte peak. Each replicate shall be independently evaluated.

$$Signal\text{-}to\text{-}Noise = \frac{height\ of\ analyte}{amplitude\ of\ noise}$$

8.7.6.3 Estimating LOD Using Statistical Analysis of Background

To determine the LOD using this approach, a minimum of three sources of blank matrix shall be analyzed in duplicate (two separate samples) over at least three runs. The average signal (e.g., integrated area of signal at the analyte's retention time) from all blank matrix samples and the standard deviation of the signal shall be calculated. Likewise, fortified matrix samples of decreasing concentration shall be analyzed in duplicate over the course of at least three runs. The lowest concentration of a fortified matrix sample that consistently yields a signal greater than the average signal from the blank matrix samples plus 3.3 times the standard deviation of the signal from the blank matrix samples, shall be identified as the LOD.

8.7.7 Estimating LOD Using a Linear Calibration Curve

This technique is useful for any quantitative method that follows a linear calibration model. A minimum of three independent calibration curves shall be constructed across the working range of the analytical method over different runs. A single source of blank matrix (per matrix type) may be used to estimate the LOD using this approach. The LOD shall be estimated from the standard deviation of the *y* intercept (s_y) and the average slope (Avg_m) as:

$$LOD = (3.3\ s_y)/Avg_m$$

8.8 Lower Limit of Quantitation

8.8.1 General

LLOQ studies shall be carried out for all quantitative methods. There are a number of different approaches for determining a method's LLOQ. Select the approach that provides the most reasonable estimation of the LLOQ given the analytical instrumentation utilized in the method. A method's LLOQ incorporates instrumental performance, as well as the sample matrix and inherent procedural limitations. Therefore, the LLOQ shall be assessed over multiple runs using fortified, matrix samples from at least three different sources per matrix type.

[8] Use of 3.3 in the LOD calculation provides a false positive error rate of 0.0005%. (see Boyd, RK, Basic C, and Bethem RA. *Trace Quantitative Analysis by Mass Spectrometry.* Hoboken, N.J.: John Wiley (2008).

8.8.2 Using the Lowest Non-Zero Calibrator as the LLOQ

In some instances, it may be sufficient to define the LLOQ as the value of the lowest non-zero calibrator. For each matrix type, at least three different blank matrix sources shall be fortified with the analyte at the concentration of the lowest calibrator and analyzed over a minimum of three runs to demonstrate that all detection, identification, bias, and precision criteria are met. If desired, it is acceptable to use the same calibrator replicates used to establish the calibration model (Section 8.3) for this approach, but additional samples/replicates may be needed to meet the minimum of nine data points including three sources per matrix type.

8.8.3 Using Decision Point Concentration as the LLOQ

In some instances, it may be sufficient to define the LLOQ as the value of an administratively-defined decision point. For example, a laboratory may choose to define a method's LLOQ for GHB as 5 mg/L for antemortem blood based on the laboratory's administratively defined decision point for reporting this analyte, even though a lower LLOQ is analytically achievable. The concentrations used for this approach shall remain within the previously established calibration curve. For each matrix type, at least three different blank matrix sources shall be fortified with the analyte at the concentration of the decision point and analyzed over a minimum of three runs to demonstrate that all detection, identification, bias, and precision criteria are met.

9 Additional Validation Parameters

9.1 General

In certain instances, it is important to evaluate additional validation parameters, if they are applicable. These include processed sample stability of the analyte(s) and the effect of sample dilution on bias and precision. A laboratory shall include these parameters in their validation plan, and determine if they are applicable to the analytical method or if they are already addressed through other means (e.g.., quality assurance practices). The laboratory validation plan shall include documentation of this evaluation.

9.2 Dilution Integrity

The effect of sample dilution shall be determined during validation of quantitative methods, if this is a routine practice within the laboratory. At times, this may be due to low specimen volume requiring the sample or assay to be adjusted appropriately. In other instances, excessively high concentrations may be encountered that are above the established calibration range. To bring the analyte concentration within the validated concentration range, the laboratory procedure may allow for reanalysis after dilution of the sample.

If dilution of a sample is allowed because of high analyte concentration or low sample volume, then the laboratory shall evaluate the effect of dilution on the method's bias and precision with at least one concentration pool. This shall be accomplished by establishing bias and within-run precision studies (as described in Section 8.2.) at a common dilution ratio (e.g., 1:2) utilized by the laboratory and determining if performance criteria are still met.

9.3 Processed Sample Stability

Processed samples are typically analyzed in batches; however, circumstances may arise in which they cannot be analyzed within a reasonable amount of time due to atypical events (e.g., instrument failures or loss of power). It may be necessary to analyze processed samples the following day or even later. In these instances, the laboratory shall evaluate the length of time processed samples can be maintained before they undergo unacceptable changes. The following approach provides a means of evaluating loss of analytes in stored, processed samples at both low and high concentrations that could impact the ability to accurately detect, identify, and quantitate them.

Processed sample stability experiments shall utilize sets of blank matrix samples fortified at both low and high concentrations. A single source of blank matrix (per matrix type) may be used to evaluate processed sample stability. The samples may be prepared by the laboratory, purchased from a commercial source, or appropriately pooled patient samples. A large enough volume of each of these samples should be used in order to complete the studies.

Numerous aliquots from each concentration set shall be processed (i.e., extracted) using the method under validation. The processed samples for a given concentration pool shall be combined, mixed well, and then divided into different (autosampler) vials for instrumental analysis. The first vials of each concentration shall be immediately analyzed in triplicate to establish the time zero responses. All remaining vials shall be maintained in a manner that they would typically be stored during routine analysis (e.g., at refrigerated or room temperature on autosampler). The remaining vials shall be analyzed in triplicate at different time intervals that represent the typical time range expected for processed samples to wait before being injected into the instrument. The analyte shall be considered stable until the average signal (e.g., peak area *or* ratios of peak area of analyte to internal standard) compared to the time zero average signal falls outside of the method's acceptable bias.

As an example, say that a method's bias limit is ±15% and the time zero average signal is 100,000. The laboratory's processed samples are placed into different autosampler vials and are analyzed repeatedly up to 72 hours. For this example, the processed sample's analyte is considered stable until the average signal falls outside of the 85,000 – 115,000 range (±15% of the time zero average signal).

A plot of the average response for each concentration pool against each time point with linear regression allows for an assessment of trends. Addendum B provides a detailed example of this approach.

10 Required Revalidation of Previously Validated Methods

Modifications to a validated method shall be evaluated to confirm that the changes do not have an adverse effect on the method's performance. The decision regarding which performance characteristics require additional validation shall be based on consideration of the specific parameters likely to be affected by the change(s). These changes may include, but are not limited to:

a) analytical conditions,

b) instrumentation,

c) sample processing,

d) data software.

For example, changes of extraction solvent or buffer may affect linearity, interferences, LLOQ, precision, and bias. A change of the analytical column stationary phase or a change in mobile phase composition may affect linearity and interferences. Further, consideration should be given to conducting parallel studies with known or proficiency samples utilizing both a previously validated method and the modified method to evaluate the effects of the changes. The goal should be to demonstrate the impact the changes have on the performance of the previously validated procedure.

11 Documentation Requirements for Method Validation

Record keeping is an essential part of laboratory operating procedures and is a key component of method validation. The data generated during method validation studies shall be maintained and available for audits, reviews, or inspections. These records shall be organized for easy retrieval and review.

Method validation records shall include a summary of the validation studies conducted and their results. The format of this summary report may be a brief bulleted report or in a table to facilitate a swift review of validation studies. The summary shall minimally include the following:

a) scope;

b) validation plan;

c) description of all the parameters evaluated, if any of the parameters were not evaluated, then the reason shall be stated or justified;

d) sample preparation steps to include concentrations and matrices;

e) raw data or reference to where the raw data are stored;

f) results and calculations;

g) conclusions;

h) references;

i) documentation of management review and approval.

The validation records shall also contain specific details regarding the studies conducted, including:

a) individuals involved in the method validation,

b) specific instrumentation,

c) dates.

Method validation documentation shall also include a copy of the newly developed analytical method or a reference to its location. Further, validation documentation should be retained for a minimum of 10 years after the method is retired.

12 Efficiency with Validation

Keep in mind that some validation experiments may be conducted concurrently with the same fortified samples. Annex C, D, and E present example approaches to assist in streamlining validation experiments.

Annex A
(informative)

Quantitation of Drug X in Blood Validation Example

The following is an example of some of the validation steps outlined in this document. It is not intended to provide specific guidance for any particular method.

In this example, assume a laboratory validated a LC/MS/MS method for a new opiate ("Drug X") in whole blood.

Create Validation Plan (Section 6)

Before starting the validation experiments, the laboratory prepared the validation plan. In the plan, they specified that an existing SPE procedure, already used for the extraction of other opiates, would be relied upon for extracting Drug X (Section 5). Further, instrument conditions were previously optimized (Section 5), so those conditions were also listed in the plan (not shown). As this is a quantitative procedure, the validation parameters listed in Table A.1—Validation Parameters to be Assessed. were assessed against the laboratory's pre-defined acceptance criteria.

Interference Studies (Section 8.5)

Ten (10) independent sources of blank whole blood were secured from previously analyzed cases to evaluate matrix interferences (Section 8.5.2). The blank matrix samples were extracted *without* the addition of internal standard (d3-Drug X) and analyzed using the newly developed method. No interferences at the retention time for Drug X were noted after analysis of the blank whole blood samples.

The laboratory randomly selected one of the blank matrix samples, added d3-Drug X to the sample (250 ng/mL), extracted the sample, and analyzed it. This was to demonstrate that the internal standard would not interfere with the signal for Drug X (Section 8.5.3). Likewise, another random blank matrix sample was fortified with Drug X at 2000 ng/mL and analyzed *without* internal standard. This was to evaluate whether the unlabeled analyte ions interfere with the signal for d3-Drug X. The results demonstrated no interferences between the analyte and internal standard.

Table A.1—Validation Parameters to be Assessed

Parameter:	Acceptance Criteria:
Bias	Shall not exceed ± 20%
Calibration Model	10 – 1000 ng/mL (linear model desired)
Carryover	Carryover after highest calibrator does not exceed 10% of signal of lowest calibrator
Interference Studies	No interfering signal from matrix, internal standard, common drugs of abuse (including other common opiates/metabolites), OTC drugs, and prescription medications
Ionization Suppression/ Enhancement	<25% suppression/enhancement and <20% CV due to matrix (if not, evaluate impact on LOD, LLOQ, and Bias)
Limit of Detection	Shall be 10 ng/mL or lower
Lower Limit of Quantitation	Shall be 10 ng/mL or lower
Precision	% CV shall not exceed 20%
Dilution Integrity	Bias and precision criteria shall be met with dilution of samples. Dilution ratios evaluated will depend on linear range of final calibration curve.
Processed Sample Stability	Evaluate length of time that analyte in extracted samples stored at room temperature on autosampler remains stable

Lastly, to evaluate interferences from other commonly encountered analytes (Section 8.5.4), the laboratory injected neat solutions diluted in mobile phase to a concentration of 5000 ng/mL (or higher) of all common opiates and metabolites observed in their casework, other common recreational drugs of abuse and their metabolites, other common prescription medications and their metabolites, and common over-the-counter medications and their metabolites. Table A.2 shows how the laboratory efficiently prepared these solutions into four injection standards. The laboratory observed no interference for the signal of Drug X or d3-Drug X from any of these compounds.

Table A.2—Example Drugs/Metabolites Used in Interference Study

Injection Mix	Included Drugs/Metabolites (5000 ng/mL unless noted otherwise)
Opiates and Related	codeine, morphine, heroin, 6-acetylmorphine, hydrocodone, hydromorphone, oxycodone, oxymorphone, levorphanol, meperidine, methadone, tramadol, fentanyl
Drugs of Abuse	amphetamine, cocaine, benzoylecgonine, ecgonine methyl ester, methamphetamine, PCP, MDA, MDMA, THC, THC-COOH
Prescription Drugs	antidepressants (amitriptyline, imipramine, doxepin, amoxapine, trazodone, bupropion, fluoxetine, sertraline, citalopram), benzodiazepines (alprazolam, chlordiazepoxide, clonazepam, clorazepate, diazepam), antiarrhythmics (verapamil, diltiazem, lidocaine), barbiturates at 500,000 ng/mL (amobarbital, butalbital, pentobarbital, phenobarbital), other CNS depressants (zopiclone, buspirone, zolpidem)
OTC Drugs	antihistamines (diphenhydramine, doxylamine, chlorpheniramine), analgesics at 500,000 ng/mL (acetaminophen, ibuprofen), antitussive (dextromethorphan)

Calibration Model (Section 8.3) and Carryover (Section 8.4)

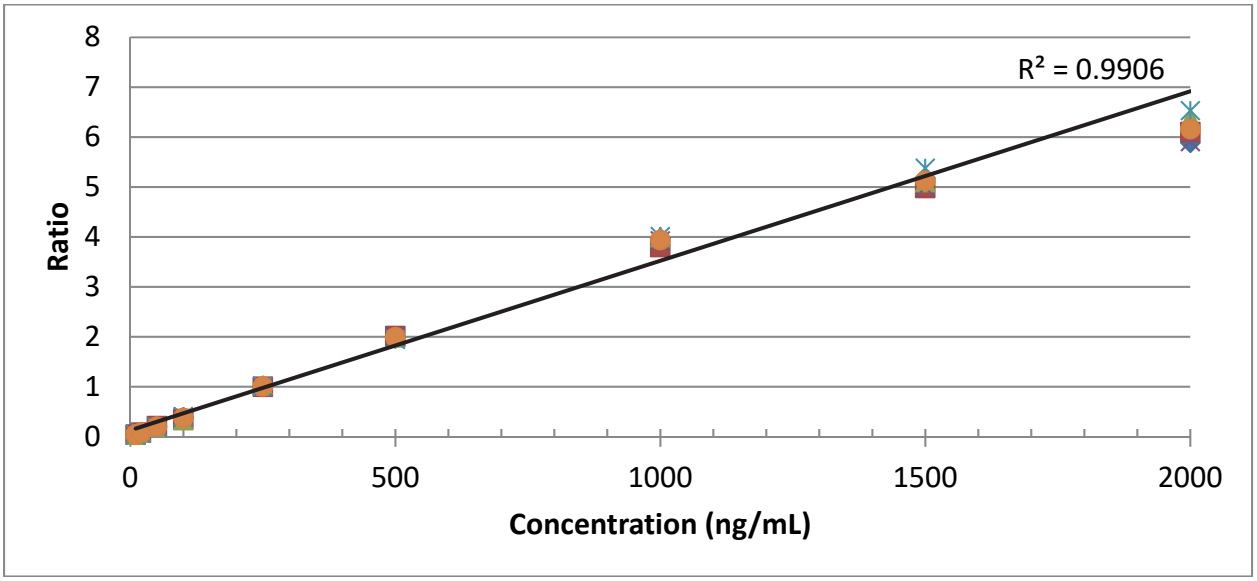
The laboratory indicated a desire for the method's calibration model to be linear and include the range of 10 – 1000 ng/mL. However, to evaluate if the method could exceed this range, the calibration samples were prepared in blank blood at the concentrations of 10, 20, 50, 100, 250, 500, 1000, 1500, and 2000 ng/mL. Each calibrator was analyzed once per run in five separate runs (Table A.3). An extracted matrix blank was analyzed after each calibrator to evaluate carryover at each concentration. The data of all runs were combined into a single calibration curve. It was noted that carryover was not present for Drug X or the internal standard in any of the extracted blank matrix samples that followed the calibrators in the range of 10 – 1500 ng/mL; however, a small amount of carryover for Drug X was observed in two of the five blank matrix samples that followed the 2000 ng/mL calibrator. The integrated areas of Drug X in these two samples were less than 10% of the smallest area of the lowest (10 ng/mL) calibrator, so the carryover from the 2000 ng/mL calibrator was deemed acceptable.

The first evaluation of the calibrator data suggested that linearity may break off above 1000 ng/mL (Table A.3 and Figure A.1). A residual plot was used to further evaluate these data (Figure A.2).

Table A.3—Calibration Curve Data

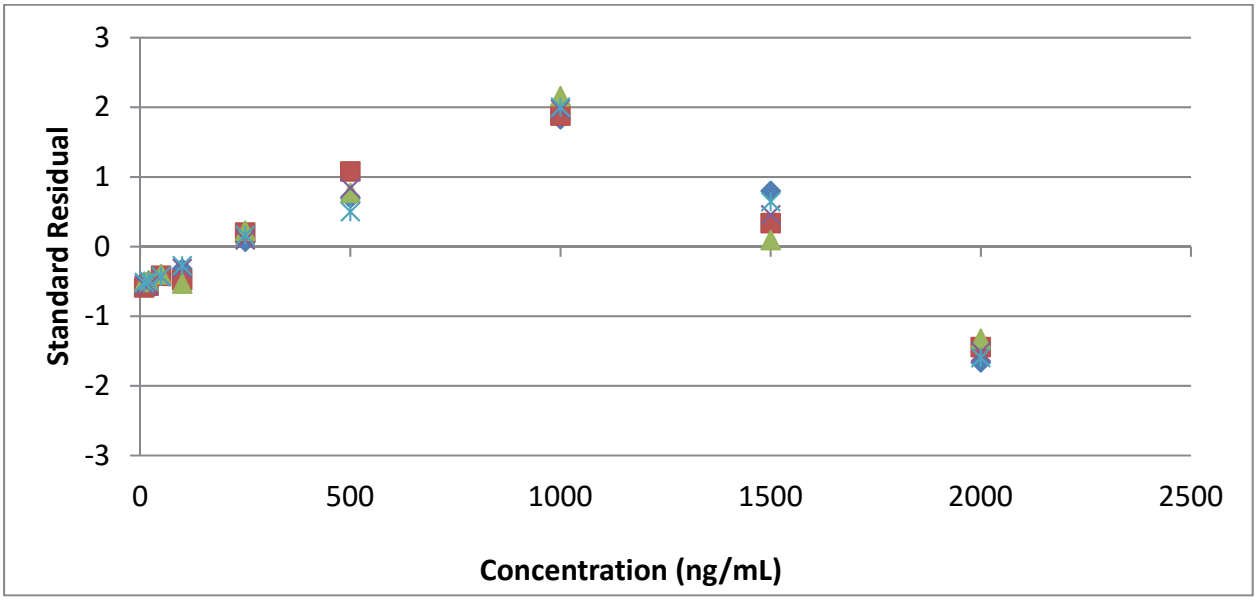
Conc (ng/mL)	Run 1			Run 2			Run 3			Run 4			Run 5		
	Peak Area		Ratio	Peak Area		Ratio	Peak Area		Ratio	Peak Area		Ratio	Peak Area		Ratio
	Drug X	Int Std		Drug X	Int Std		Drug X	Int Std		Drug X	Int Std		Drug X	Int Std	
10	3951	101310	**0.039**	4112	100281	**0.041**	3971	101833	**0.039**	4319	102831	**0.042**	3872	101885	**0.038**
20	7937	105831	**0.075**	7930	100382	**0.079**	8254	105821	**0.078**	7511	104321	**0.072**	8154	103215	**0.079**
50	20470	100838	**0.203**	21753	103588	**0.21**	20302	102018	**0.199**	21590	105832	**0.204**	20266	103927	**0.195**
100	39703	102328	**0.388**	38273	108729	**0.352**	34141	103771	**0.329**	40784	102731	**0.397**	41763	103887	**0.402**
250	100584	100887	**0.997**	108547	108439	**1.001**	105062	102800	**1.022**	100227	100832	**0.994**	107173	105382	**1.017**
500	211080	105382	**2.003**	215146	106772	**2.015**	207431	104394	**1.987**	208572	103819	**2.009**	203887	104665	**1.948**
1000	412751	103889	**3.973**	379832	99982	**3.799**	417110	104382	**3.996**	395587	100838	**3.923**	403819	100728	**4.009**
1500	517537	99872	**5.182**	477045	95773	**4.981**	483159	94737	**5.1**	475616	95333	**4.989**	525150	97666	**5.377**
2000	560931	95283	**5.887**	561994	92100	**6.102**	593173	93076	**6.373**	531467	89942	**5.909**	606638	92886	**6.531**

Figure A.1—Combined Calibration Curve Demonstrating Loss of Linearity Above 1000 G/ML



The residual plot showed an inverted U-shaped distribution suggesting a non-linear model would be the best calibration model for these data (Figure A.2).

Figure A.2—Standard Residual Plot of Calibration Curve Data with an Inverted U-Shaped Distribution



Because the laboratory's preference was to use a linear calibration model, they re-evaluated these data after dropping the 1500 and 2000 ng/mL calibrators. Doing so allowed for their original validation plan requirements (10 ng/mL – 1000 ng/mL) to still be met. The revised calibration curve appeared to provide a better fit of these data using an unweighted linear model (Figure A.3). This was confirmed by the residual plot that showed a random distribution around the zero line suggesting a linear model was the most appropriate for these data (Figure A.4).

Figure A.3—Revised Calibration Curve

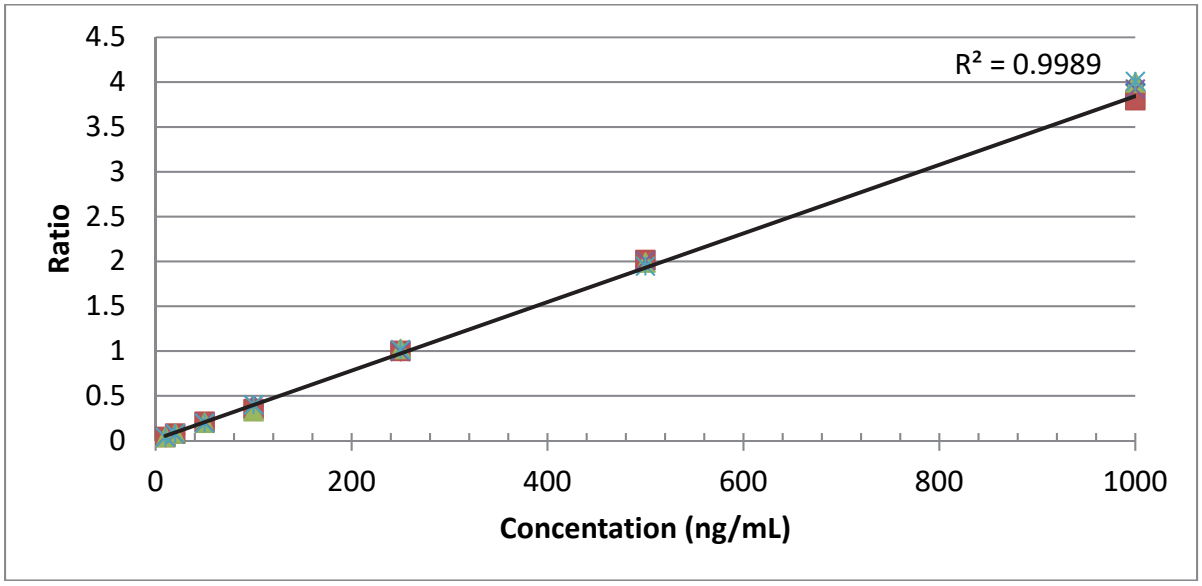
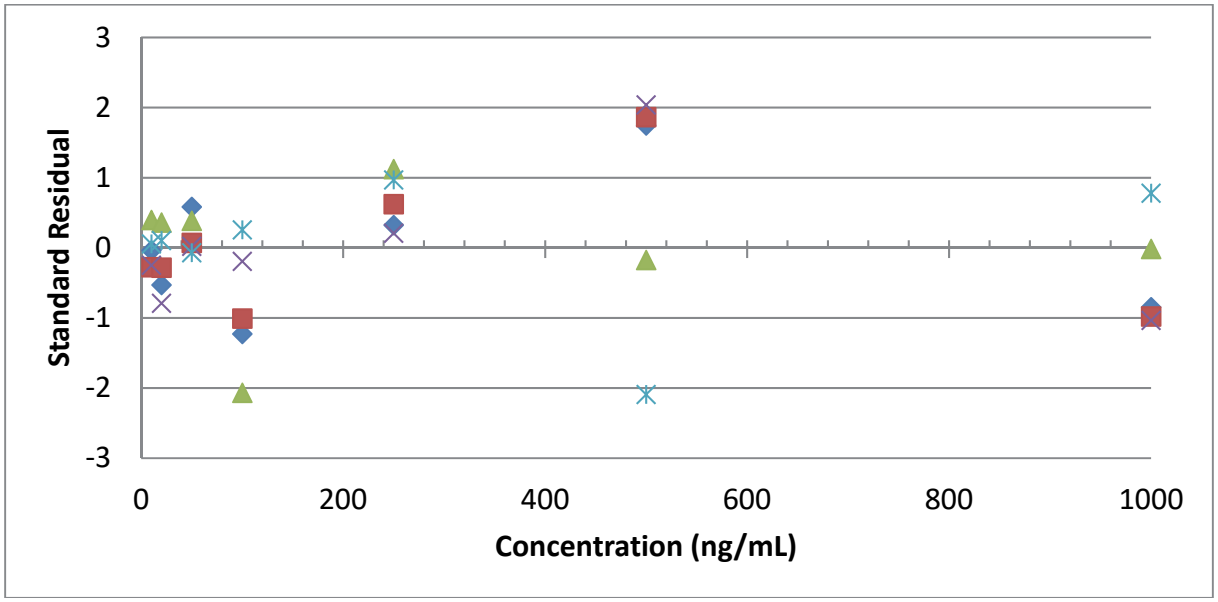


Figure A.4—Residual Plot of Calibration Curved Data with a Random Distribution



For future validation experiments, the laboratory used calibrators prepared at 10, 50, 100, 250, 500, and 1000 ng/mL.

Since accurate quantitative results cannot be assumed above 1000 ng/mL, the laboratory knew they would have to re-extract (with dilution) any samples that exceed 1000 ng/mL. Therefore, they planned to evaluate dilution integrity in ratios up to 1:5 when conducting the bias and precision experiments.

Recall that no significant carryover was observed up to 2000 ng/mL in the laboratory's study. However, since the working calibration range will end at 1000 ng/mL, the laboratory recognized that accurate quantitative results cannot be achieved above the working range. So the laboratory defines how they will address carryover by evaluating all samples that immediately follow those that exceed 1000 ng/mL. If the amount of Drug X in these samples is above the method's LOD, the samples with potential carryover will be re-extracted and analyzed.

LOD (Section 8.7)

To estimate the LOD, the laboratory chose to utilize the results from their previously generated calibration curve data (Section 8.7.7). Both the slope and *y*-intercept of the individual calibration curves (10 – 1000 ng/mL) were determined in order to calculate the average slope and standard deviation of the *y*-intercept s (Table A.4).

Table A.4—Slope and Y-Intercept Data from Calibration Curves

	Slope	*y*-intercept
Run 1	0.003980	-0.00050
Run 2	0.003828	-0.01543
Run 3	0.004009	-0.01247
Run 4	0.003934	0.00695
Run 5	0.003995	-0.00318
Average	**0.003949**	0.00125
Std Dev	0.000073	**0.01054**

The LOD was calculated using the formula: *3.3 × 0.01054 / 0.003949 = 8.8 ng/mL*

LLOQ (Section 8.8)

The laboratory chose to use the lowest calibrator as their LLOQ (Section 8.8.2.). Three sources of whole blood were fortified at 10 ng/mL. The samples were each extracted and analyzed against a freshly prepared calibration curve on three different runs for a total of nine data points. The results demonstrated that this concentration was capable of reproducibly providing symmetrical peaks and the minimum mass spectral identification ratios, while maintaining a bias of ±20% and a % CV of <20%. Therefore, 10 ng/mL was confirmed as the method's LLOQ.

Bias and Precision (Section 8.2)

To establish the method's bias and precision, the laboratory prepared three pools of fortified matrix samples at the following concentrations: low (30 ng/mL); medium (400 ng/mL); and high (800 ng/mL). Each concentration pool was analyzed in triplicate within five separate runs along with a freshly prepared calibration curve (Table A.5).

The laboratory calculated the bias (Section 8.2.1) by first determining the mean for each concentration. This resulted in the values listed in Table A.6.

From these values, the bias was calculated at each concentration. For example, for the low concentration sample, the bias was determined as: $Bias_{Low} = ((28 – 30)/30) \times 100) = (- 6.7\%)$

Likewise, the bias for the medium and high concentrations was calculated as 9.3% and –2.4%, respectively.

Table A.5—Quantitative Results (ng/mL) of Bias and Precision Runs

Low (30 ng/mL)	Run 1	Run 2	Run 3	Run 4	Run 5
Rep 1	32	26	29	26	28
Rep 2	28	24	31	35	25
Rep 3	27	28	27	30	29
Med (400 ng/mL)	**Run 1**	**Run 2**	**Run 3**	**Run 4**	**Run 5**
Rep 1	412	435	427	455	444
Rep 2	444	410	419	438	442
Rep 3	422	450	479	452	423
High (800 ng/mL)	**Run 1**	**Run 2**	**Run 3**	**Run 4**	**Run 5**
Rep 1	892	793	761	742	820
Rep 2	827	741	729	734	749
Rep 3	850	769	803	720	791

Table A.6—Mean Concentrations (ng/mL) for Bias Calculations

Conc (ng/mL)	Calculated Mean	Bias
Low (30)	28	-6.7%
Med (400)	437	9.3%
High (800)	781	-2.4%

Within-run and between-run precisions were calculated using the one-way ANOVA approach (Section 8.2.2.3.4). Using the ANOVA: Single Factor analysis in a spreadsheet program (see Table A.7 for Low Concentration), the laboratory was able to obtain the value for the mean square within groups (MSwg) for the low concentration and introduced it into the appropriate formula as follows:

$$Within-run\ CV(\%) = \left[\frac{\sqrt{MS_{wg}}}{grand\ mean\ for\ each\ concentration}\right] \times 100$$

$$Within-run\ CV(\%) = \left[\frac{\sqrt{7.933}}{28}\right] \times 100$$

$$Within-run\ CV(\%) = \left[\frac{2.817}{28}\right] \times 100$$

$$Within-run\ CV(\%) = 10.1\%$$

Table A.7—ANOVA Calculations for 30 ng/mL Sample

ANOVA: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Column 1	3	87	29	7
Column 2	3	78	26	4
Column 3	3	87	29	4
Column 4	3	91	30.33333	20.33333
Column 5	3	82	27.33333	4.333333

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	34	4	8.5	1.071429	0.420175	3.47805
Within Groups	79.33333	10	7.933333			
Total	113.3333	14				

The between-run precision for the low concentration was calculated using the formula and the mean square between groups (MSbg) from the ANOVA table:

$$Between - Run\ CV(\%) = \left[\frac{\sqrt{\frac{MS_{bg} + (n-1) * MS_{wg}}{n}}}{grand\ mean\ for\ each\ concentration}\right] \times 100$$

$$Between - Run\ CV(\%) = \left[\frac{\sqrt{\frac{8.5 + (3-1) * 7.933}{3}}}{28}\right] \times 100$$

$$Between - Run\ CV(\%) = \left[\frac{2.845}{28}\right] \times 100$$

$$Between - Run\ CV(\%) = 10.2\%$$

Using the data for the medium and high concentrations, the ANOVA: Single Factor analysis was also conducted on these concentration pools (data not shown) and appropriate values introduced into the formulas to obtain the within-run and between-run precisions. Table A.8 lists the calculated results for all concentrations.

Table A.8—Precision Results

	Low	Medium	High
Within-Run	10.1% CV	4.5% CV	3.9% CV
Between-Run	10.2% CV	4.2% CV	2.2% CV

Ionization Suppression/Enhancement (Section 8.6)

As the instrumental portion of the method involves LC/MS/MS, the laboratory needed to conduct ionization suppression/enhancement experiments. The post-column extraction approach was chosen (Section 8.6.3).

Two sets of samples were prepared for the experiment. Set one consisted of standards prepared in mobile phase at 30 and 800 ng/mL. They were not extracted, but instead simply injected six times each.

Set two was prepared in ten blank matrix samples. Each blank matrix was from an independent source of blank whole blood from previously analyzed cases. These were the same ten blank matrix samples used in the interference studies. Each blank matrix sample was extracted in duplicate *and then* fortified to 30 and 800 ng/mL with Drug X and 250 ng/mL with d3-Drug X. Each concentration set sample was injected one time each.

Average peak areas for both the Drug X and the d3-Drug X are found in Table A.9.

Table A.9—Average Peak Areas From Suppression/Enhancement Experiments

Average Peak Areas				
	30 ng/mL		800 ng/mL	
	Drug X	d3-Drug X	Drug X	d3-Drug X
Set 1	13890	110381	330822	112827
Set 2	11812	102444	303992	105923

Using the above data sets, the laboratory calculated the % ionization suppression/ enhancement for the target ion transitions at each concentration using the formula:

$$\%\ Ionization\ suppression/enhancement_{Drug\ X\ (Low)} = ((11812 / 13890)-1) \times 100 = (-15.0\%)$$

The negative value suggested some suppression was occurring, but it was less than 25%.

Similarly, the laboratory calculated the % suppression/enhancement for 800 ng/mL and for the internal standard in both sets. The results suggested suppression of -8.1% for Drug X at the 800 ng/mL concentration. Although at the same concentration in both the low and high samples, the d3-Drug X demonstrated ionization suppressions of 7.2% and 6.1%, respectively.

The data were also used to calculate the % CV at each concentration. All % CVs were <14% (data not shown).

Since the average suppression for all analytes did not exceed ±25% and the calculated % CV value was <20%, the variation was considered insignificant. No further work was required for other validation parameters.

Dilution Integrity (Section 9.2)

While the laboratory indicated that a minimum working range for the calibration curve was between 10 – 1000 ng/mL, they anticipated occasional samples containing Drug X would exceed concentrations above 1000 ng/mL. Their initial attempt to extend the calibration range to 2000 ng/mL was abandoned when they realized that a non-linear calibration model would be needed. Therefore, they conducted dilution integrity experiments to demonstrate acceptable bias and precision results when samples are diluted in deionized water. They evaluated a dilution ratio of 1:5.

The laboratory prepared a fortified matrix sample at a concentration of 3000 ng/mL. The 3000 ng/mL sample was diluted 1:5 and analyzed in triplicate over five different runs; each with a freshly prepared calibration curve. Bias and within-run precision calculations were performed and results (Table A.10) demonstrated comparable values compared to the results obtained without dilution. This provided proof of no detrimental impact when diluting the samples before extraction.

Table A.10—Effect of Dilution on Bias and Precision

	3000 ng/mL (1:5 dilution)
Bias	9.9%
Within-Run Precision	2.9%

Processed Sample Stability (Section 9.3)

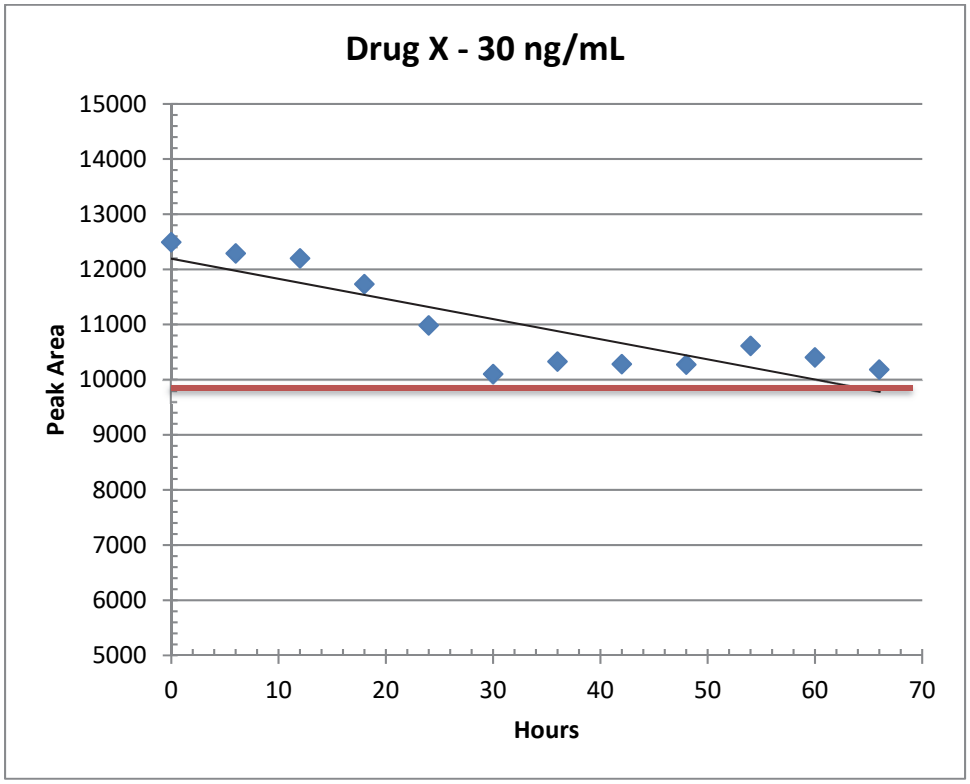
The laboratory recognized that samples are not always analyzed immediately after extraction due to large batches or unforeseen delays. For example, the instrument may lose communication with its controller, inadvertently stopping the analytical run. Therefore, to evaluate the impact of room temperature storage of processed samples sitting in queue on the autosampler before analysis, the laboratory conducted a stability study on extracted samples. This was achieved by preparing fortified matrix samples at two concentrations, 30 ng/mL and 800 ng/mL. Twelve aliquots of each concentration were extracted. Reconstituted extracts for each concentration were combined and vortexed to ensure adequate mixing. The concentration pool was then divided into 12 autosampler vials and placed on the autosampler. The first vial of each level was injected three times to represent the time zero (t_0) sample. The remaining vials for each concentration were analyzed in triplicate every six hours up to 66 hours. Analyte signals from the triplicate analyses were averaged and compared to the t_0 signals (Table A.11).

Table A.11—Average Peak Areas for Processed Sample Stability Study

Time (hr)	Average Peak Area			
	30 ng/mL		800 ng/mL	
	Drug X	d3-Drug X	Drug X	d3-Drug X
0	12490	101832	332554	100423
6	12289	100382	331820	100328
12	12198	100432	330779	101101
18	11732	100733	330246	100987
24	10983	100992	329787	100832
30	10101	101789	326048	100821
36	10328	100904	327238	100234
42	10281	100086	326838	100323
48	10271	100183	315009	99727
54	10612	100309	315772	99421
60	10402	100233	316231	96381
66	10183	100872	315499	94832

By plotting the average peak areas for both Drug X and the internal standard, the laboratory could evaluate the processed samples while they were stored on the autosampler. As their required bias is ±20%, they considered the compounds stable until they saw a decrease (or increase) in signal of more than 20% from the t_0 average peak area. The plot for the 30 ng/mL concentration of Drug X is shown to demonstrate this concept (Figure A.5).

Figure A.5—Change in Drug X Peak Area Over 66 Hours



These data appear to suggest Drug X remained stable within the pre-defined limits for the entire 66-hour period of the study. However, the trend line shows that 66 hours may be the maximum period of time before the samples may need to be re-extracted. It was noted that at the 30-hour mark, stability seemed to have dropped very close to the "instability" point. Since the previously determined bias was actually much better than the ±20% required in their validation plan, the laboratory made a decision to re-extract any samples that remain on the autosampler more than 24 hours.

Documentation of Results (Section 11)

Along with all of the other required documentation listed in Section 11, the laboratory compared the results from their studies to their originally defined requirements, as demonstrated in Table A.12. This summary was evidence that the method was fit-for-purpose and thereby valid to analyze whole blood for Drug X.

Table A.12—Summary of Validation Results

Parameter:	**Acceptance Criteria:**	**Result:**
Bias	Shall not exceed ± 20%	-6.7 to 9.3%
Calibration Model	10 – 1000 ng/mL (linear model desired)	10 – 1000 ng/mL (linear model)
Carryover	Carryover after highest calibrator does not exceed 10% of signal of lowest calibrator.	No significant carryover at 2000 ng/mL. Re-extract and analyze samples containing Drug X above the LOD if that sample follows one that exceeds 1000 ng/mL of Drug X.
Interference Studies	No interfering signal from matrix, internal standard, common drugs of abuse (including other common opiates/metabolites), OTC drugs, and prescription medications.	No observed interferences from matrix or from common drugs/metabolites.
Ionization Suppression/ Enhancement	<25% suppression/enhancement and <20% CV due to matrix (if not, evaluate impact on LOD, LLOQ, and Bias).	-8.1 to -15.0%; <14% CV
Limit of Detection	Shall be 10 ng/mL or lower	8.8 ng/mL
Lower Limit of Quantitation	Shall be 10 ng/mL or lower	10 ng/mL
Precision	% CV shall not exceed 20%	Within-run (3.9 to 10.1%) Between-run (2.2 to 10.2%)
Dilution Integrity	Bias and precision criteria shall be met with dilution of samples.	Using a 1:5 aqueous dilution, bias (9.9%) and within-run precision (2.9%). Comparable to results obtained without dilution.
Processed Sample Stability	Evaluate length of time that analyte in extracted samples stored at room temperature remains stable	24 hours

Annex B
(informative)

Immunoassay Screen of Benzodiazepines in Urine Validation Example

The following is an example of the immunoassay validation steps outlined in this document. It is not intended to provide specific guidance for any particular method.

In this example, assume a laboratory validated an immunoassay kit for its ability to screen urine for select benzodiazepines.

Create Validation Plan (Section 6)

Before starting the validation experiments, the laboratory prepared the validation plan. In the plan, they specified that they will use Company ABC's ELISA Immunoassay Kit for Benzodiazepines (Oxazepam) designed with a "cutoff" of 300 ng/mL. The laboratory planned to select their own cutoff concentration (decision point) of 50 ng/mL for the target compound of oxazepam. The validation parameters were assessed against the pre-defined requirements listed in Table B.1.

Table B.1—Validation Parameters to be Assessed

Parameter:	Acceptance Criteria:
Limit of Detection	Oxazepam (target): 50 ng/mL Nordiazepam: 50 ng/mL Lorazepam: 100 ng/mL Alprazolam: 25 ng/mL Alpha-Hydroxyalprazolam: 50 ng/mL
Precision	% CV shall not exceed 20%; grand means ± 2 Std Dev of low and high concentration pools will not overlap with grand mean of decision point

Limit of Detection (Section 8.7.2) and Precision at the Decision Point [Cutoff] (Section 8.2.2.2)

The laboratory wanted to use a lower cutoff concentration (50ng/mL) than that established by the manufacturer (300 ng/mL), so they were required to evaluate the precision of the immunoassay at their new cut-off concentration (instead of at the manufacturer-set cutoff concentration). The laboratory prepared three pools of oxazepam-fortified matrix samples at the following concentrations: 25 ng/mL (50% below); 50 ng/mL (decision point); and 100 ng/mL (100% above). Each of the fortified sample sets was analyzed in triplicate within five separate runs. A blank matrix sample was also analyzed in each run to establish B_0. The results are shown in Table B.2.

The grand mean result for the 50 ng/mL decision point concentration was 30.444%. The two standard deviation ranges calculated for the samples at 25 ng/mL (36.803% – 46.575%) or 100 ng/mL (15.952% – 24.560%) did not overlap with the mean value for the decision point concentration.

The % CV for the 25, 50, and 100 ng/mL concentration pools were 5.9%, 5.1%, and 10.6%, respectively; well below the requirement to not exceed 20%.

Table B.2—Results (B/B_0 for ELISA) of Precision Runs for the Oxazepam Sample Sets

25ng/mL	Run 1	Run 2	Run 3	Run 4	Run 5
Rep 1	44.4%	39.9%	39.2%	41.2%	41.9%
Rep 2	43.9%	40.6%	37. 6%	39.1%	41.3%
Rep 3	45.3%	44.8%	39. 8%	41.5%	44.8%
Grand Mean			**41.7%**		
Std Dev			**2.4%**		
% CV			**5.9%**		
50 ng/mL	**Run 1**	**Run 2**	**Run 3**	**Run 4**	**Run 5**
Rep 1	32.7%	31.9%	30.6%	29.8%	32.7%
Rep 2	29.9%	30.9%	28.2%	31.0%	32.3%
Rep 3	27.7%	31.1%	29.6%	28.9%	29.3%
Grand Mean			**30.4%**		
Std Dev			**1.6%**		
% CV			**5.1%**		
100 ng/mL	**Run 1**	**Run 2**	**Run 3**	**Run 4**	**Run 5**
Rep 1	19.3%	24.0%	20.9%	23.3%	23.3%
Rep 2	17.0%	18.9%	19.5%	20.2%	20.8%
Rep 3	17.8%	18.7%	17.9%	20.2%	22.0%
Grand Mean			**20.3%**		
Std Dev			**2.2%**		
% CV			**10.6%**		

The product brochure listed the cross-reactivities for the select benzodiazepines as follows: oxazepam (100%); nordiazepam (425%); lorazepam (50%); alprazolam (450%); and alpha-hydroxyalprazolam (340%).

For this example, the laboratory wanted to inform their customers (using this immunoassay kit) that they can screen for nordiazpam and alpha-hydroxyalprazolam at the same concentration as the target (oxazepam) target cut-off concentration (50 ng/mL). Since the product brochure indicates the cross-reactivities for these two analytes are greater than 100%, the laboratory did not need to conduct any further validation for these two analytes using this company's immunoassay kit, *because they have demonstrated the kit performs within the precision limits for oxazepam when the cut-off concentration was lowered to 50 ng/mL.*

In contrast, the laboratory mathematically estimated the detection limit for lorazepam to be 100 ng/mL (due to 50% cross-reactivity) when oxazepam is used as the target analyte at 50 ng/mL. Since they have lowered the manufacturer's original cutoff concentration for the target analyte (oxazepam), they also need to confirm that the 100 ng/mL LOD for lorazepam achieves adequate precision because it has poorer cross-reactity than oxazepam. Precison experiments, similar to those conducted for oxazepam above, were conducted for lorazepam at the 100 ng/mL decision point, as well as concentrations -50% and +100% of the decision point (data not shown).

Additionally, the laboratory wanted to report to their customers their capability of detecting alprazolam at 25 ng/mL. Even though alprazolam cross-reacts much better than oxazepam according to the product brochure, the laboratory's desire to use a detection limit *below* that of the target analyte requires validation that the immunoassay kit can appropriately detect alprazolam at the lower concentration with adequate precision. Precison experiments, similar to those conducted for oxazepam above, were conducted for alprazolam at the 25 ng/mL decision point, as well as concentrations -50% and +100% of the decision point (data not shown).

Documentation of Results (Section 11)

Along with all of the other required documentation listed in Section 11, the laboratory compared the results from the validation studies conducted to the originally defined requirements, as demonstrated in Table B.3. This summary was evidence that the immunoassay method, as modified, was fit-for-purpose and thereby valid to analyze urine for oxazepam, lorazepam, alprazolam, and alpha-hydroxyalprazolam.

Table B.3—Summary of Validation Results

Parameter:	Acceptance Criteria:	Result:
Limit of Detection	Oxazepam (50 ng/mL) Nordiazepam (50 ng/mL) Lorazepam (100 ng/mL) Alprazolam (25 ng/mL) Alpha-hydroxyalprazolam (50 ng/mL)	50 ng/mL 50 ng/mL 100 ng/mL 25 ng/mL 50 ng/mL
Precision	% CV shall not exceed 20%; grand means ± 2 Std Dev of low and high concentration pools will not overlap with grand mean of decision point	Oxazepam: % CV: 5.1 to 10.6% 25 ng/mL: 36.8 – 46.6% 50 ng/mL: 30.4% 100 ng/mL: 16.0 – 24.6% Lorazepam (data not shown): % CV: 9.8 to 13.8% 50 ng/mL: 39.8 – 48.1% 100 ng/mL: 30.4% 200 ng/mL: 16.9 – 25.0% Alprazolam (data not shown): % CV: 6.9 to 11.9% 13 ng/mL: 25.9 – 32.3% 25 ng/mL: 17.1% 50 ng/mL: 8.2 – 10. 7% NOTE: Nordiazepam and alpha-hydroxyalprazolam cross react better than the target analyte (oxazepam) and their desired LODs are at the same concentration. So no further evaluation of precision is required for these analytes.

Annex C
(informative)

Table of Example Experiments for Validation of Qualitative Confirmation/Identification Methods[9]

Interference (Section 8.5)
— 10 different sources of each matrix, no IS (Section 8.5.2) — 1 blank sample + isotopically-labeled IS (Section 8.5.3) — 1 fortified sample with high analyte concentrations, no IS (Section 8.5.3) — Neat, fortified, or authentic samples containing potentially interfering compounds/metabolites but not the analyte(s) of interest (Section 8.5.4)
Carryover (Section 8.4)
— Addressed in routine QC practices by analyzing extracted blank matrix samples between case samples
Limit of Detection (Section 8.7.6.2)
— Reference Material Approach (Section 8.7.6.2.): Blank matrix samples from three different sources of each matrix type fortified at increasingly lower concentrations and analyzed in duplicate over 3 runs. Lowest concentration in each matrix that reproducibly yields signal greater than or equal to 3.3 times the noise of background signal AND achieves the predefined detection and identification criteria is considered as the LOD for that matrix.

[9] LC/MS methods also require ionization suppression/enhancement experiments (Section 8.6)

Annex D
(informative)

Table of Example Experiments for Validation of LC/MS –Based Quantitative Methods

Interference (Section 8.5)
— 10 different sources of each matrix, no IS (Section 8.5.2) — 1 blank sample + isotopically-labeled IS (Section 8.5.3) — 1 fortified sample with high analyte concentrations, no IS (Section 8.5.3) — Neat, fortified, or authentic samples containing potentially interfering compounds/metabolites but not the analyte(s) of interest (Section 8.5.4)
Calibration Model (Section 8.3)
— 6 concentration levels. 5 replicates of each in separate runs. Data combined.
Ionization Suppression/Enhancement (Section 8.6)
— Post-Extraction Addition Approach (8.6.3.): 2 sets of 10 matrix-matched, blank sample extracts fortified at low and high concentrations after extraction. Neat standards (low and high) injected 6 times.
Carryover (Section 8.4)
— Addressed in routine QC practices by analyzing extracted blank matrix samples between case samples.
Bias and Precision (Section 8.2)
— 3 concentration levels, triplicate analyses of each concentration in each of 5 separate runs with new calibration curve for each run (Sections 8.2.1. and 8.2.2.3.)
Limit of Detection (Section 8.7.6.2)
— Decision Point Approach (Section 8.7.5.): Three different sources of blank matrix samples fortified at decision point concentration and analyzed over 3 runs to demonstrate acceptable detection and identification criteria are met.
Limit of Quantitation (Section 8.8)
— Decision Point Approach (Section 8.8.3.): Three different sources of blank matrix samples fortified at decision point concentration and analyzed over 3 runs to demonstrate acceptable bias and precision criteria are met.

Annex E
(informative)

Table of Example Experiments for Validation of Quantitative Methods

Interferences	Ionization Suppression/Enhancement [a]	Calibration Model
—10 different sources of each matrix, without IS —1 blank sample with IS —1 fortified sample with high analyte concentrations and without IS —Neat, fortified, or authentic samples containing potentially interfering compounds/metabolites but no analyte	**Post-column infusion:** • 10 blank extracts fortified after extraction at low concentration • 10 blank extracts fortified after extraction at high concentration • Analyte solutions for infusion (low and high concentrations) each injected 6 times	• 6 concentration levels, 5 replicates each (may be accomplished with calibration curves generated for studies below)

Main validation phase							
		Bias & Precision					*Dilution Integrity*
Run	*Calibration*	*Low*	*Medium*	*High*	*LOD*[b]	*LLOQ*[b]	*Bias & Precision*
1	6	3	3	3	3	3	3
2	6	3	3	3	3	3	3
3	6	3	3	3	3	3	3
4	6	3	3	3	-	-	3
5	6	3	3	3	-	-	3

[a] LC/MS(/MS) methods only
[b] For this example, the lowest non-zero calibrator approach is used to estimate the LOD (Section 8.7.4) and LLOQ (Section 8.8.2.)

NOTE Carryover addressed in routine QC practices by analyzing extracted blank matrix samples between case samples.

Annex F
(informative)

Bibliography

1] Araujo, P. "Key Aspects of Analytical Method Validation and Linearity Evaluation." *Journal of Chromatogrraphy B,* Vol. 877, 2009, pp. 2224-2234.

2] Boyd, R.K., Basic, C. and Bethem, R.A. *Trace Quantitative Analysis by Mass Spectrometry.* Hoboken, N.J., John Wiley, 2008.

3] Bressolle, F., Bromet-Petit, M. and Audran, M. "Validation of Liquid Chromatographic and Gas Chromatographic Methods, Applications to Pharmacokinetics." *Journal of Chromatogrraphy B,* Vol. 686, 1996, pp. 3-10.

4] Bruce, P., Minkkinen and Riekkola, M. "Practical Method Validation: Validation Sufficient for an Analytical Method." *Mikrochim Acta*, Vol. 128, 1998, pp. 93-106.

5] Clinical and Laboratory Standards Institute. *C50-A Mass Spectrometry in the Clinical Laboratory: General Principles and Guidance; Approved Guideline.* C50- A, Vol 27, No 24, 2007.

6] Corley, J. "Best Practices in Establishing Detection and Quantification Limits for Pesticides Residues in Foods." *Handbook of Residue Analytical Methods for Agrochemicals*, John Wiley & Sons Ltd, Chichester, 2003.

7] Drummer, O. "Requirements for Bioanalytical Procedures in Postmortem Toxicology." *Analytical and Bioanalytical Chemistry,* Vol. 388, 2007, pp. 1495-1503.

8] Eurachem Guide. *The Fitness for Purposes of Analytical Methods.* 1998.

9] Health Sciences Authority. *Guidance Notes on Analytical Method Validation: Methodology.* 2004.

10] Hubert, Ph., Nguyen-Huu, J-J., Boulanger. B., et al. "Harmonization of Strategies for the Validation of Quantitative Analytical Procedures ASFSTP Proposal – Part I." *Journal of Pharmaceutical and Biomedical Analysis*, Vol. 36, 2004, pp. 579-586.

11] Hubert, Ph., Nguyen-Huu, J-J., Boulanger. B., et al. "Harmonization of Strategies for the Validation of Quantitative Analytical Procedures ASFSTP Proposal – Part II." *Journal of Pharmaceutical and Biomedical Analysis,* Vol. 45, 2007, pp. 70-81.

12] Hubert, Ph., Nguyen-Huu, J-J., Boulanger, B., et al. "Harmonization of Strategies for the Validation of Quantitative Analytical Procedures ASFSTP Proposal – Part III." *Journal of Pharmaceutical and Biomedical Analysis,* Vol.45, 2007, pp, 82-96.

13] Kushnir, M.M., Rockwood, A. L., Nelson, G. J., et al. "Assessing Analytical Specificity in Quantitative Analysis Using Tandem Mass Spectrometry.", *Clinical Biochemistry*, Vol. 12:003, 2004, pp. 319-327.

14] Liang, H. R., Foltz, R. L., Meng, M., Bennett, P. "Ionization enhancement in atmospheric pressure chemical ionization and suppression in electrospray ionization between target drugs and stable-isotope-labeled internal standards in quantitative liquid chromatography/tandem mass spectrometry." *Rapid Communications in Mass Spectometry*, Vol. 17, 2003, pp. 2815-2821.

15] Matuszewski, B. K., Constanzer, M. L., Chavez-Eng, C. M. "Strategies for the Assessment of Matrix Effects in Quantitative Bioanalytical Methods Based on HPLC-MS/MS." *Analytical Chemistry*, Vol. 75:13, 2003, pp. 3019-3030.

16] Niedbala, R. S. and Gonzalez, J.M. "Immunoassays." *Clarke's Analysis of Drugs and Poisons,* 4th edition, edited by Moffat, A.C., Osselton, M.D., Widdop, B., and Watts, J., Pharmaceutical Press, London, 2011.

17] ISO/IEC Guide 99:2007, *International Vocabulary of Metrology – Basic and General Concepts and Associated Terms*. VIM, Organization for Standardization (ISO)/International Electrotechnical Commission (IEC). Geneva, 2007.

18] Peters, F., Maurer, H. H. "Bioanalytical Method Validation and Its Implications for Forensic and Clinical Toxicology – A Review." *Accreditation and Quality Assurance,* Vol, 7, 2002, pp. 441-449.

19] Peters, F.T. "Method Validation." *Applications of LC-MS in Toxicology*, Pharmaceutical Press, London, 2006.

20] Peters, F., Drummer, O., Musshoff, F. "Validation of New Methods." *Forensic Science International*, Vol. 165, 2007, pp. 216-224.

21] Remane, D., Meyer, M.R., Wissenbach, D.K., Mauer, H.H. "Ion suppression and enhancement effects of co-eluting analytes in multi-analyte approaches: systematic investigation using ultra-high performance liquid chromatography/mass spectrometry with atmospheric pressure chemical ionization or electrospray ionization." *Rapid Communications in Mass Spectometry*, Vol, 24, 2010, pp. 3103-3108.

22] Shah, V., Midha, K., Dighe, S., et al. "Analytical Methods Validation: Bioavailability, Bioequivalence and Pharmacokinetics Studies." *Pharmaceutical Research*, Vol. 9:4, 1992, pp. 588-592.

23] Shah, V., Midha, K., Findlay, J., et al. "Bioanalytical Method Validation – A Revisit with a Decade of Progress." *Pharmaceutical Research,* Vol. 17:12, 2000, pp. 1551-1557.

24] Shultz, E.K. "Analytical Goals and Clinical Interpretation of Laboratory Procedures" *Tietz Textbook of Clinical Chemistry,* 2nd edition, Burtis, C.A and Ashwood, E.R., Eds. W.B. Saunders Company, Pennsylvania, 1994.

25] Stockl, D., D'Hondt, H., Thienpont, L. M. "Method Validation Across the Disciplines- Critical Investigation of Major Validation Criteria and Associated Experimental Protocols." *Journal of Chromatography B*, Vol. 877, 2009, pp. 2180-2190.

26] Thompson, M., Ellison, S. L. R., Wood, R. "Harmonized Guidelines for Single-Laboratory Validation of Methods of Analysis, IUPAC Technical Report." *Pure and Applied Chemistry*, Vol. 74:5, 2002, pp. 835-855.

27] U.S. Department of Health and Human Services, Food and Drug Administration. *Guidance for Industry, Bioanalytical Method Validation*. 2001.

28] Van Eeckhart, A., Lanckmas, K., Sarre, S., et al. "Validation of Bioanalytical LC-MS/MS Assays: Evaluation of Matrix Effects." *Journal of Chromatography B*, Vol. 877, 2009, pp. 2198-2207.

29] Viswanathan, C. T., Bansal, S., Booth. B., et al. "Quantitative Bioanalytical Methods Validation and Implementation: Best Practices for Chromatographic and Ligand Binding Assays." *Pharmaceutical Research*, Vol. 24:10, 2007, pp. 1962-1973.

30] Viswanathan, C. T., Bansal, S., Booth, B., et al. "Workshop/Conference Report - Quantitative Bioanalytical Methods Validation and Implementation: Best Practices for Chromatographic and Ligand Binding Assays." *The AAPS Journal,* Vol. 9:1, 2007, pp. E30-E42.

Academy Standards Board
410 North 21st Street
Colorado Springs, CO 80904

www.asbstandardsboard.org